

RESEARCH ARTICLE

How well do we know ourselves? Disentangling self-judgment biases in perceived accuracy and preference of personality feedback

SABINA RAMONA TRIF

Department of Psychology, Work and Organizational Psychology Research Center, Babeş-Bolyai University, Cluj-Napoca, Romania

CLAUDIA RUS

Department of Psychology, Work and Organizational Psychology Research Center, Babeş-Bolyai University, Cluj-Napoca, Romania

ELENA MANOLE

Department of Psychology, Work and Organizational Psychology Research Center, Babeş-Bolyai University, Cluj-Napoca, Romania

OCTAVIAN CĂLIN DUMA

Faculty of Business, Babeş-Bolyai University, Cluj-Napoca, Romania

Abstract

Despite personality measurement and feedback being pervasive practices, there are self-judgment biases that may impair their usage. We set out to analyze the differences between two kinds of false feedback and real feedback on personality regarding perceived accuracy and preference. We propose that there would be no differences between false and real feedback regarding perceived accuracy, but we expect differences regarding feedback preference. A sample of 146 students completed the IPIP-50 instrument that measured the Big 5 Factors and received three kinds of feedback - a general one (Barnum effect as false feedback), a positive one (Better-than-average effect as false feedback), and a real one. They rated each regarding accuracy and preference. Results indicate differences regarding both dependent variables. Participants perceive false feedback as more accurate than the real one. Moreover, they prefer positive feedback over the other two, and general feedback compared to the real one. We discuss both theoretical and practical implications, alongside a series of limitations and future research directions.

Keywords

personality, Barnum effect, better-than-average effect, psychometrics

Introduction

Personality measurement is a pervasive element in organizations, especially regarding employee selection (Nikolaou & Foti, 2018),

as it predicts multiple work-relevant outcomes, such as engagement (Young et al., 2018), creativity and voice behaviors (Zare & Flinchbaugh, 2019) and performance (Zell & Lesick, 2022). Yet, the majority of

* Correspondence concerning this article should be addressed to Sabina Ramona Trif, Department of Psychology, Work and Organizational Psychology Research Center, Babeş-Bolyai University, 37 Republicii Street, 400015, Cluj-Napoca, Romania. Email: sabinatrif@psychology.ro

participants rate bogus or trivial personality statements as accurate and personally applicable, a phenomenon known as the Barnum effect (for reviews, see Dickson & Kelly, 1985; Furnham & Schofield, 1987; Snyder et al., 1977).

Previous studies have investigated whether the content of feedback statements plays a role in ratings of accuracy and acceptance (Furnham & Varian, 1988; Halperin et al., 1976; Macdonald & Standing, 2002). Because feedback was more frequently accepted when using generally positive statements over neutral or negative ones, researchers argued that the Barnum effect could overlap with the Better-Than-Average effect (BTA; Johnson et al., 1985; MacDonald & Standing, 2002). This effect explains the tendency to view ourselves in a positive light and is supported by studies highlighting only a moderate association between self-evaluations of ability and objective ratings of performance (Dunning et al., 2004).

There has been a lack of explicit distinction between ratings of perceived accuracy and preferences in previous work on the Barnum effect. Researchers also frequently only compared Barnum effect feedback to either real feedback or feedback presenting the individual in a positive light (Andersen & Nordvik, 2002; Dana & Fouke, 1979; MacDonald & Standing, 2002). Additionally, the literature to date shows contradictory findings regarding the existence of the Barnum effect, although the majority of studies report results in favor of this bias. This limits our understanding of biases in the judgment of personality feedback. Such biases are crucial in understanding how well people generally know or understand themselves. Individuals are highly likely to endorse global, ambiguous, or positive descriptions and integrate them into their self-concept, especially if they believe they were derived from the results of psychological assessment techniques, and specifically developed for them (Johnson et al., 1985; Snyder et al., 1977).

In order to disentangle the literature on measurement acceptance and preference, we set out to investigate whether participants differentiate between generalized, better-than-average, and real, personalized personality

feedback in terms of accuracy. Secondly, drawing on previous research on biases in self-judgments, we test whether participants prefer a generalized (Barnum effect) profile over a real one, or a profile presenting them in a positive light (better-than-average effect).

The present study contributes to the literature on self-judgment biases by differentiating between perceived accuracy and preference in the context of three types of feedback. Consistent with the evidence supporting the Barnum effect, we argue that individuals are biased in their judgment of personality feedback. Moreover, we argue that participants will more likely endorse and prefer personality feedback that presents them in a self-enhancing way.

Theoretical framework

Self-knowledge and personality assessment

Personality is most frequently conceptualized in terms of traits, defined as relatively enduring patterns of thought, emotion, and behavior (Funder, 2001; McCrae & John, 1992). The Big Five model (also known as the Five-Factor model or FFM) is one of the most widely used personality taxonomies (Costa & McCrae, 1992; Goldberg, 1993). This model defines five broad superordinate dimensions (extraversion, agreeableness, conscientiousness, neuroticism, and openness or intellect), capturing individual differences in personality (McCrae & Costa, 1987; John & Srivastava, 1999).

Personality assessment consists of procedures for determining and measuring the underlying characteristics that account for individual differences in personality structure (Ozer & Reise, 1994). During the reporting stage, the examiner and the assessee frequently engage in a collaborative assessment process, when the examiner presents their feedback and the assessee can contribute with their own view (Fischer, 2000). This process is especially relevant when examiners are less certain about their conclusions, or when the assessment is undertaken during a validation study. Validating an instrument based on individuals' appraisal of accuracy is based on the underlying assumption that individuals can

accurately assess their personality or other intrinsic traits, often better than an observer would. For example, Costa and McCrae (1992) used ratings of the accuracy of NEO-PI-R profiles to assess the validity of their instrument.

However, drawing on the Barnum effect literature, participants should be more likely to endorse assessments and accept feedback as accurate if the statements used are generally accurate (trivial) rather than specific. On the other hand, drawing on self-enhancement theories, participants should be more likely to endorse positive statements, irrespective of their actual accuracy (Dickson & Kelly, 1985).

This led authors to question the validity of accuracy ratings and the practice of using participants' own assessment of instruments as a validation criterion (Forer, 1949), as participants might readily accept a profile as personally relevant and assimilate that content into their own self-concept (Rogers & Soule, 2009). Similarly, Guerrettaz and Arkin (2016) argue that individuals who possess a high degree of self-concept clarity rate Barnum statements as less accurate, while individuals with lower self-concept clarity can more easily endorse the non-specific feedback and integrate it into their self-knowledge.

Biases in the accuracy of self-judgment

The Barnum effect

The Barnum effect (also known as the Forer effect; Forer, 1949) characterizes people's tendency to accept bogus personality feedback as accurate descriptions of their personality. Although this effect was first introduced by Forer (1949), it was later labeled as the Barnum effect by Meehl (1956), as this scholar thought psychology clients were agreeing to general profile descriptions that could apply to everyone. Scholars have extensively investigated this effect, demonstrating the strong tendency to accept vague and generalized feedback after completing a personality test (Dana & Fouke, 1979; Holmes et al., 1986; Rogers & Soule, 2009; Stachnik & Stachnik, 1980).

Although Barnum descriptions are general in nature and applicable to the general

population, participants frequently rate Barnum-type personality feedback as personally relevant and uniquely true about themselves. This prompted many researchers to conclude that participants were gullible and lacked self-knowledge, considering generalized reports significantly more applicable to themselves than others (Johnson et al., 1985; Standing & Keays, 1987). This argument has been used to explain why many individuals believe in astrology and fortune-tellers (Dickson & Kelly, 1985; Fichten & Sunerton, 1983; Snyder, 1974). However, there are other critical perspectives of the Barnum effect that go against the gullibility hypothesis.

One perspective is that statements are accepted because they are vague in nature, which means participants can internalize the traits described based on their own understanding and assessment of the reports (Dickson & Kelly, 1985; Merrens & Richards, 1970).

Another perspective holds that statements are accepted because they are generally true, so they accurately describe the majority of individuals. Accordingly, participants are rational in rating the statements as true of themselves; in this case, the statements are considered to have high "base-rate accuracy" (Layne, 1979; Michels & Layne, 1980; Verberne, 1974). Especially when asked directly, participants can recognize that Barnum effect feedback fails to describe their uniqueness (Greene, 1977) and can differentiate between accurate, trivial, and inaccurate feedback (Harris & Greene, 1984).

Putting the two perspectives together, the observed differences could be due to methodological choices. Generally, participants discern more easily between real and randomly generated personality feedback compared to real personality feedback and Barnum-effect feedback (Andersen & Nordvik, 2002; Christian & Bringmann, 1982; De Fruyt & Wille, 2013). For example, participants were able to correctly identify descriptions based on randomly generated personality scores that deviated from their original scores (Andersen & Nordvik, 2002). Also, these participants tended to disagree more with the false feedback reports when the

difference between the real and random scores was larger (De Fruyt & Wille, 2013). However, participants still endorsed false reports if they were based on neutral scores (De Fruyt & Wille, 2013).

Consistent with the previously cited literature, we put forward the following hypothesis:

Hypothesis 1: The perceived accuracy of the real personality assessment report will be similar to the perceived accuracy of the false personality assessment reports including the Barnum effect feedback and the better-than-average effect feedback.

Although participants rate both real and fake personality descriptions as accurate, previous research suggests differences in terms of preference. Specifically, there is evidence that participants favor generalized (i.e., Barnum effect) feedback over personalized, real feedback that is based on their actual test scores. When asked to choose which profile better describes their personality, participants favored the Barnum-effect description over a real description based on their actual Personality Research Form scores (Merrens & Richards, 1970). Accordingly, we put forward the following hypothesis:

Hypothesis 2: The participants will accept the Barnum effect feedback more frequently than in the case of the real feedback on personality assessment.

The better-than-average effect

Literature reveals that participants are more likely to favor descriptions containing positive statements or traits over negative statements, consistent with the better-than-average effect (Michels & Layne, 1980). When asked to rate the accuracy of astrological descriptions, even non-believers rated the fake astrological profiles as accurate when they contained positive statements (Glick et al., 1989). When directly comparing Barnum descriptions, participants frequently rate positively worded descriptions as more accurate and more descriptive of their own personalities than negatively worded descriptions (Furnham & Varian, 1988; Halperin et al., 1976; Macdonald & Standing, 2002).

The better-than-average effect describes people's tendency to evaluate their abilities, attributes, and personality traits as better than an average peer (Alicke & Govorun, 2005). This effect comes in line with the self-enhancement bias as self-insight (Allport, 1960; Kwan et al., 2004), where individuals perceive themselves in a better light than others do. In support of the self-enhancement bias, this effect is more likely to be observed for positive traits, showing that individuals are more prone to present and view themselves in a positive light rather than self-protect and minimize their negative characteristics. Moreover, this effect is larger for personality traits than abilities (Zell et al., 2020).

Therefore, we formulate the following hypothesis:

Hypothesis 3: The participants will accept the favorable (BTA) feedback more frequently than in the case of the real feedback on personality assessment (3a) and the Barnum effect feedback (3b).

Method

Participants

The sample comprises 146 participants that were enrolled as bachelor students in a large public Romanian university. Most of the participants were female (90.40%). The average age of the participants was 20.56 years old ($SD = 1.55$). The sample of students has not previously received any personality feedback during the course they were attending. However, we do not have collected any information on previous personality assessments they have received.

Participation in this study was voluntary, although all the tasks were carried out as part of a class on psychology applied to work and organization. The completion of the IPIP-50 instrument was mandatory for the final grade of the participants, yet participation in the feedback evaluation was not mandatory. Those that took part in the feedback evaluation task received extra points for their final grade.

Instruments

To measure personality we used the Romanian version of the International Personality Item Pool 50 (IPIP-50) (Iliescu et al., 2015). Ten items were used to measure each of the following personality factors: extraversion ($\alpha = .89$; "I am the life of the party"),

agreeability ($a = .80$; “*I am interested in people*”), conscientiousness ($a = .69$; the exclusion of a reversed item leads to an alpha of $.71$; “*I pay attention to details*”), emotional stability ($a = .83$; “*I get stressed out easily*”), and openness ($a = .79$; “*I have excellent ideas*”). Participants were asked to rate each item using a Likert scale, ranging from 1 (Totally disagree) to 5 (Totally agree).

We examined two dependent variables – perceived global perceived accuracy of the description (which is defined as the measure of perceived global accuracy for the descriptions) and description acceptance (which is defined as a measure of evaluating each one of the statements that compose a description).

To assess the global perceived accuracy of the personality descriptions we used a single-item measure that was originally used by Forer (1949) - “*On a scale from 0 to 5, where 0 represents “completely disagree” and 5 represents “completely agree”, please assess to what extent this description encaptures your personality*”. This scale was used in the case of all three types of personality profiles/feedback: real feedback, Barnum effect, and BTA effect.

For description acceptance, the second dependent variable, we used a single-item scale inspired by Forer (1949). Participants were asked to re-read the personality descriptions and then rate whether each paragraph is true, false, or unsure. Each of the alternatives had an assigned value - the “true” answer was coded with 1 point, a “false” answer was coded with minus 1 and an “unsure” was coded with 0. In order to compute the final score, we counted the number of “true” answers and divided it by the total number of statements. In order to work with integers, we multiplied the scores by 100.

Procedure

After the participants gave their consent to be part of this study, they were asked to fill in the Romanian version of the IPIP-50 (Iliescu et al., 2015). Each student that filled in this instrument, received an electronic report including three personality profiles: (a). the real personality profile resulted from IPIP-50;

(b). a fake personality profile meant to position the individual as above average including five paragraphs developed by Paterson (in Forer, 1949; e.g. “*Above average in intelligence or mental alertness. Also above average in accuracy—rather painstaking at times. Deserves a reputation for neatness—dislikes turning out sloppy work. Has initiative; that is, ability to make suggestions and to get new ideas, open-mindedness*”), and (c). a fake personality profile that included 13 general short statements reflecting universal descriptions, used by Forer (1949) (e.g. “*You prefer a certain amount of change and variety and become dissatisfied when hemmed in by restrictions and limitations*”). Participants were asked to assess the perceived validity of the descriptions and the feedback acceptance.

In order to avoid the order effect, we split the participants into two sub-groups. A number of the 78 participants were included in one of the sub-groups and 68 participants were included in the second one. The first sub-group received the fake feedback first, while the second group received the true feedback first. Data was collected through Google forms between April 2020-May 2021.

Results

Data was analyzed using IBM SPSS V26. In order to test the first hypothesis, regarding the similarity in the perceived accuracy between real and fictitious descriptions, we used non-parametric tests. The choice of the non-parametric tests was driven by the measurement of the dependent variable on an ordinal scale. We chose the Friedman test as we adopted a repeated-measures design that involves an independent variable with three levels (i.e., better-than-average feedback, Barnum feedback, and real feedback on personality). In order to test the second hypothesis, regarding the acceptance of feedback, we used the repeated-measures analysis of variance (ANOVA).

Table 1 presents the means, standard deviations, and correlation coefficients for the personality measures, perceived accuracy, feedback acceptance, and gender.

According to the first hypothesis, we expect similar levels of perceived accuracy for

the three types of feedback (i.e., better-than-average, Barnum, and real feedback). In order to test this hypothesis we used the Friedman non-parametric test. Data does not support this hypothesis, as we identified a significant difference between the three kinds of feedback ($\chi^2(2) = 36.19, p < .001$).

As the Friedman test is an omnibus test, we were further interested in identifying the specific types of feedback where significant differences emerge. Thus, we used the Wilcoxon test in order to make two-by-two comparisons. In order to avoid the escalation of the alpha error, we used the Bonferroni

correction. We divided the .05 threshold by the number of comparisons, resulting in a threshold of .016. The comparison between the two false feedbacks (i.e., Barnum and better-than-average feedback) does not show a statistically significant difference ($z = -2.12, p = .03$). The comparison between the real feedback and better-than-average feedback shows a statistically significant difference ($z = -5.41, p < .001$). The comparison between the real feedback and the Barnum feedback also shows a statistically significant difference ($z = -4.12, p < .001$). Table 2 presents the rankings for the comparisons.

Table 1. Correlations, means, and standard deviations ($N = 146$)

Variable	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11
1. Gender													
2. Extraversion	3.30	.81	-.02	-									
3. Agreeability	3.99	.56	.16	.32**									
4. Conscientiousness	3.67	.54	.09	.02	.14								
5. Emotional Stability	2.83	.72	-.10	.18*	-.01	.31**							
6. Intellect	3.66	.53	.03	.42**	.20*	-.03	-.08						
7. Accuracy BTA	4.07	.93	.20*	.27**	.24**	.13	.02	.09					
8. Accuracy Barnum	3.90	.88	.18*	.06	.05	-.04	-.29**	.09	.44**				
9. Accuracy Real	3.46	1.07	.05	-.01	.01	.06	.08	-.01	.25**	.31**			
10. Acceptance BTA	.86	.18	.23*	.39**	.19*	.04	.12	.13	.64**	.25**	.11		
11. Acceptance Barnum	.72	.19	.15	.10	.11	-.17*	-.38**	.21*	.33**	.58**	.18*	.26**	
12. Acceptance Real	.61	.25	.12	.00	.02	.03	-.06	.02	.25**	.29**	.69**	.12	.27**

Note. On the diagonal line we preset the means and standard deviations for each variable; * $p < .05$, ** $p < .01$

According to the second and third hypotheses, there are significant differences regarding the acceptance of feedback, with participants accepting the Barnum feedback more than the real one (H2), and the better-than-average feedback more than the other two types of feedback (i.e., general and real) (H3). In order to test these hypotheses we used repeated measures ANOVA, with gender as a control variable.

Overall, there is a significant difference regarding the acceptance of feedback ($F(2, 288) = 19.60, p < .001, \eta^2 = .12, \pi = 1$). The effect of gender is also significant ($F(1, 144) = 8.55, p < .01, \eta^2 = .056, \pi = .828$), with females ($M = .74, SE = .01$) accepting feedback at higher levels compared to males ($M = .63, SE = .04$). Table 3 shows the pairwise comparisons, where we used the Bonferroni test. The second hypothesis is

supported by data, as the average acceptance for the Barnum feedback ($M = 68.07$, $SE = 2.61$) is significantly higher than the one for the real feedback ($M = 56.62$, $SE = 3.56$). The third hypothesis also received empirical

support, as the average acceptance for the better-than-average feedback ($M = 80.07$, $SE = 2.50$) is significantly higher than the other two kinds of feedback.

Table 2. *Wilcoxon test ranks*

		<i>N</i>	Mean Ranks	Sum of Ranks
Barnum and Better-than-Average Feedbacks	Negative ranks	49	38.60	1891.50
	Positive ranks	28	39.70	1111.50
	Ties	69		
	Total	146		
Real and Better-than-Average Feedbacks	Negative ranks	74	51.62	3820
	Positive ranks	23	40.57	933
	Ties	49		
	Total	146		
Real and Barnum Feedbacks	Negative ranks	63	45.48	2865
	Positive ranks	24	40.13	963
	Ties	59		
	Total	146		

Table 3. *Pairwise comparisons feedback acceptance (N = 146)*

Variable 1	Variable 2	Mean Difference	<i>SE</i>	95% Confidence Interval
Better-Than-Average Feedback	Barnum Feedback	11.99***	3.16	[4.30, 19.65]
Better-Than-Average Feedback	Real Feedback	23.44***	4.16	[13.37, 33.51]
Barnum Feedback	Real Feedback	11.45**	3.84	[2.14, 20.76]

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Discussion

The aim of the present paper was twofold. First, we were interested in differences in the perceived feedback accuracy regarding personality assessment for false and real feedback. Second, we were interested in differences in the preference for false and real feedback on personality profiles

Overall, our results show support for the existence of biases in the judgment of personality feedback. Primarily, our first set of

data analyses provide findings that were not in line with the expectations included in our first hypothesis. These findings indicate a tendency to rate fictitious feedback as more accurate than real one. This comes in support of the Barnum effect (Dickson & Kelly, 1985; Furnham & Schofield, 1987; Snyder et al., 1977), as more general descriptions of personality are easily integrated into the self-concept of participants (Dickson & Kelly, 1985; Merrens & Richards, 1970).

Interestingly, the better-than-average effect feedback (Alicke & Govorun, 2005) was also assessed by our participants as more accurate than the real feedback, with no differences between the two types of false feedback. These results may offer a more clear direction toward unraveling the mechanism for these biases. The results regarding perceived accuracy may also be explained through the lens of the self-enhancement effect (Dickson & Kelly, 1985). Thus, in order for feedback to enable one to gain higher self-esteem levels and to provide a positive self-image, it must be accepted as accurately describing the person. This can be achieved with either general, easy-to-integrate statements (i.e. Barnum effect), or overly positive ones (i.e. better-than-average effect).

Secondarily, our participants not only rated fictitious feedback as more accurate, but they also preferred it over real feedback. This preference indicates a significantly higher acceptance in the case of Barnum effect feedback and better-than-average feedback descriptions. These findings are in line with the previous research on the Barnum effect showing a preference for general, rather than actual feedback (Merrens & Richards, 1970). This is also aligned with the self-enhancement effect (Dickson & Kelly, 1985), with participants preferring feedback that enables them to maintain a positive self-image. There was also a significant difference in preference between the two false feedbacks, with participants preferring the better-than-average feedback over the Barnum effect. This difference underscores the idea that people are more likely to endorse feedback that puts them in a good light when it comes to personality.

Theoretical and practical implications

Our findings have multiple implications, both theoretically and practically. In terms of theoretical implications, we add to and extend the literature on the Barnum effect and the better-than-average effect, showing support for the idea that individuals are more open to feedback that is easy to integrate (i.e. Barnum Effect) or that puts them in a positive light (i.e., Better-than-Average Effect). Furthermore, apart from previous research, we examined

simultaneously these three types of feedback on personality descriptions. Interestingly, from the perspective of accuracy, there was not a significant difference between the two types of false feedback. But there was a significant difference regarding the preference for the three types of feedback. These findings take the study of self-judgment biases a step forwards, as it becomes clear that people are biased in their views on feedback, but the new research venue involves disentangling the subtle differences between biases and the relationship between them.

Additionally, our findings discuss self-judgment biases in light of two dependent variables - perceived accuracy and preference, clearly delineating between the biased perception of feedback and the preference towards a positively-formulated one. We note that participants do not only prefer more general and positive feedback, but they also perceive it as being more accurate. These results are important as they offer insight into the pervasive influences of feedback in the light of the self-enhancement effect (Dickson & Kelly, 1985), with general and positive feedback being more readily accepted due to their effect on a positive self-concept.

At the practical level, these results have important implications regarding the usage of personality measurements and feedback, in general, and when working with student populations in academic contexts, in particular. In these contexts, personality assessment may be used for self-knowledge and career orientation purposes. As participants readily tend to accept general and positive feedback, it would be worthwhile to provide real feedback on personality by presenting information in a positive manner (e.g., positively worded, discussing positive consequences) in order to ensure their integration. Moreover, as the theory backing these results involves the self-enhancement effect, with people being motivated to maintain a positive self-image (Dickson & Kelly, 1985), practitioners could build the feedback around the positive self-concept of the assessee, helping bridge the gaps between the positive self-image and the real results of the assessment. Thus, they could identify the positive self-concept of the assessee, and their current position, and discuss ways in which they can attain that image.

Limitations and future directions

The present paper also has limitations. First, the sample of our study consists of students derived from one bachelor specialization from one Romanian University. As the participation in this study was voluntary, this limits the generalizability of the results to the general population, in general, and to the student population, in particular. Moreover, it could be that age may place a role in the observed results. It is possible that our findings hold only for the population of younger people. As we collected data from university students, our study was insensitive to group differences in terms of age effects. Research suggests that senior students were more skeptical than juniors or sophomores in rating Barnum effects statements (Beins, 1993; Greene, 1977), so we encourage the replication of our study using samples of older students or other categories of participants that are assessed in terms of personality for various decisions, such as employment or career-related decisions.

Another limitation regards the content of the task we used, particularly how the three types of feedback on personality were formulated. As the personality assessment targeted a five-factor personality model (Costa & McCrae, 1992; Goldberg, 1993), only the real feedback was constructed around the five dimensions. The two kinds of false feedback were extracted from studies regarding self-perception biases (Forer, 1949), and they had patterns of response that were not reflective of the five-factor personality model (Costa & McCrae, 1992; Goldberg, 1993). As research showed that the degree of discrepancy between different types of feedback in terms of scaffolding or format plays an important role (De Fruyt & Wille, 2013), false feedback should be constructed by respecting the same patterns and carefully manipulating the quantity of discrepancy. Moreover, it is possible that the differences we identified between the three types of feedback in terms of perceived accuracy and preference might reflect differences in the language used in the content of the personality descriptions such as jargon. Literature reveals that psychological reports often involve jargon that imposes a high level of reading difficulty, which can become problematic in that they are read by

multiple audiences with varying levels of educational background (Harvey, 2006) and experience with psychological testing and assessment, in general, and personality assessments, in particular.

Furthermore, we consider that the lack of high-stake feedback consequences may limit the generalizability of the present findings. In organizations, personality feedback usually has some consequences, such as employment decisions (Nikolaou & Foti, 2018), performance improvement, and general development (Jelley, 2021; Kwiatkowski, 2018). In the present study, there were no high-stake situations regarding the personality assessment results, which may not have led the participants to the need to protect the self as much. Future studies may focus on high-stake situations, where rejection, exposure of weak points, or penalization may trigger the need to protect the self-image.

Last, future research could include the manipulation of assessment and feedback reports length. In this sense, previous studies indicate that longer assessment procedures, but shorter personality descriptions are more often endorsed than shorter assessment procedures, respectively longer personality descriptions (Dickson & Kelly, 1985).

Conclusion

The present study represents an effort to disentangle the effect of different types of feedback on judgments of accuracy and preference. We used three types of feedback, two reflecting self-judgment biases (i.e., Barnum effect, better-than-average effect), and one reflecting the observed scores on a personality assessment instrument. We used a within-subjects experimental design, in which participants rated the accuracy and preference for the three types of feedback on personality. Our findings revealed that participants rate false feedback as more accurate than real feedback. In addition, participants prefer better-than-average feedback more than the other two kinds. They also prefer the general, Barnum feedback more than the real one. These results add to the literature on self-judgment biases and offer insights for practitioners regarding personality assessment in general, and reporting, in particular.

References

- Alicke, M. D., & Govorun, O. (2005). The Better-Than-Average Effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The Self in Social Judgment* (pp. 85–106). Psychology Press.
- Allport, G. W. (1960). *Personality and social encounter: Selected essays*. Oxford, England: Beacon.
- Andersen, P., & Nordvik, H. (2002). Possible Barnum Effect in the Five Factor Model: Do Respondents Accept Random Neo Personality Inventory–Revised Scores as Their Actual Trait Profile? *Psychological Reports*, 90(2), 539–545. <https://doi.org/10.2466/pr0.2002.90.2.539>
- Beins, B. C. (1993). Using the Barnum Effect to Teach About Ethics and Deception in Research. *Teaching of Psychology*, 20(1), 33–35. https://doi.org/10.1207/s15328023top2001_6
- Christian, J. K., & Bringmann, W. G. (1982). Comparison of Computerized versus Standardized Feedback and Accurate versus Inaccurate Feedback. *Psychological Reports*, 50(3_suppl), 1067–1070. <https://doi.org/10.2466/pr0.1982.50.3c.1067>
- Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment*, 4(1), 5–13. <https://doi.org/10.1037/1040-3590.4.1.5>
- Dana, R. H., & Fouke, H. P. (1979). Barnum statements in reports of psychological assessment. *Psychological Reports*, 44(3_suppl), 1215–1221. <https://doi.org/10.2466/pr0.1979.44.3c.1215>
- De Fruyt, F., & Wille, B. (2013). “Hey, This is not like me!” Convergent validity and personal validation of computerized personality reports. *European Review of Applied Psychology*, 63(5), 287–294. <https://doi.org/10.1016/j.erap.2013.07.001>
- Dickson, D. H., & Kelly, I. W. (1985). The ‘Barnum Effect’ in Personality Assessment: A Review of the Literature. *Psychological Reports*, 57(2), 367–382. <https://doi.org/10.2466/pr0.1985.57.2.367>
- Downey, J. L. (1999). Sex and the Barnum Effect: A Reply to Layne. *Psychological Reports*, 84(2), 424–426. <https://doi.org/10.2466/pr0.1999.84.2.424>
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological science in the public interest*, 5(3), 69–106. <https://doi.org/10.1111/j.1529-1006.2004.00018.x>
- Fichten, C. S., & Sunerton, B. (1983). Popular Horoscopes and the “Barnum Effect.” *The Journal of Psychology*, 114(1), 123–134. <https://doi.org/10.1080/00223980.1983.9915405>
- Fischer, C. T. (2000). Collaborative, individualized assessment. *Journal of Personality Assessment*, 74(1), 2–14. <https://doi.org/10.1207/S15327752JPA740102>
- Forer, B. R. (1949). The fallacy of personal validation: a classroom demonstration of gullibility. *The Journal of Abnormal and Social Psychology*, 44(1), 118–123. <https://doi.org/10.1037/h0059240>
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197–221. <https://doi.org/10.1146/annurev.psych.52.1.197>
- Furnham, A., & Schofield, S. (1987). Accepting personality test feedback: A review of the Barnum effect. *Current Psychology*, 6(2), 162–178. <https://doi.org/10.1007/BF02686623>
- Furnham, A., & Varian, C. (1988). Predicting and accepting personality test scores. *Personality and Individual Differences*, 9(4), 735–748. [https://doi.org/10.1016/0191-8869\(88\)90063-3](https://doi.org/10.1016/0191-8869(88)90063-3)
- Glick, P., Gottesman, D., & Jolton, J. (1989). The Fault is not in the Stars: Susceptibility of Skeptics and Believers in Astrology to the Barnum Effect. *Personality and Social Psychology Bulletin*, 15(4), 572–583. <https://doi.org/10.1177/0146167289154010>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Greene, R. L. (1977). Student acceptance of generalized personality interpretations: A reexamination. *Journal of Consulting and Clinical Psychology*, 45(5), 965–966. <https://doi.org/10.1037/0022-006X.45.5.965>
- Guerretaz, J., & Arkin, R. M. (2016). Distinguishing the subjective and the objective aspects of self-concept clarity. *Social and Personality Psychology Compass*, 10(4), 219–230. <https://doi.org/10.1111/spc3.12243>
- Halperin, K., Snyder, C. R., Shenkel, R. J., & Houston, B. K. (1976). Effects of source status and message favorability on acceptance of personality feedback. *Journal of Applied Psychology*, 61, 85–88. <https://doi.org/10.1037/0021-9010.61.1.85>
- Harris, M. E., & Greene, R. L. (1984). Students’ Perception of Actual, Trivial, and Inaccurate Personality Feedback. *Journal of Personality Assessment*, 48(2), 179–184. https://doi.org/10.1207/s15327752jpa4802_13
- Harvey, V. S. (2005). Variables affecting the clarity of psychological reports. *Journal of Clinical Psychology*, 62(1), 5–18. <https://doi.org/10.1002/jclp.20196>
- Holmes, C. B., Buchannan, J. A., Dungan, D. S., & Reed, T. (1986). The Barnum effect in Luscher color test interpretation. *Journal of Clinical Psychology*, 42(1), 133–136. [https://doi.org/10.1002/1097-4679\(198601\)42:1%3C133::AID-JCLP2270420122%3E3.0.CO;2-7](https://doi.org/10.1002/1097-4679(198601)42:1%3C133::AID-JCLP2270420122%3E3.0.CO;2-7)
- Iliescu, D., Popa, M., & Dimache, R. (2019). Adaptarea românească a Setului Internațional de Itemi de Personalitate: IPIP-Ro. *Psihologia Resurselor Umane*, 13(1), 83–112.
- Jelley, R. B. (2021). Using personality feedback for work-related development and performance improvement: A rapid evidence assessment. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, 53(2), 175. <https://psycnet.apa.org/doi/10.1037/cbs0000230>
- John, O. P., & Srivastava, S. (1999). The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (pp. 102–138). Guilford Press.
- Johnson, J. T., Cain, L. M., Falke, T. L., Hayman, J., & Perillo, E. (1985). The “Barnum effect” revisited: Cognitive and motivational factors in the acceptance of personality descriptions. *Journal of Personality and Social Psychology*, 49(5), 1378–1391. <https://doi.org/10.1037/0022-3514.49.5.1378>
- Kwiatkowski, R. (2018). Inside-out and outside-in: The use of personality and 360 degree data in executive coaching. In *Executive Coaching* (pp. 153–181). Routledge.
- Kwan, V. S. Y., John, O. P., Kenny, D. A., Bond, M. H., & Robins, R. W. (2004). Reconceptualizing Individual Differences in Self-Enhancement Bias: An

- Interpersonal Approach. *Psychological Review*, 111(1), 94–110. <https://doi.org/10.1037/0033-295X.111.1.94>
- Layne, C. (1979). The Barnum effect: Rationality versus gullibility? *Journal of Consulting and Clinical Psychology*, 47, 219–221. <https://doi.org/10.1037/0022-006X.47.1.219>
- Layne, C. (1998). Gender and the Barnum Effect: A Reinterpretation of Piper-Terry and Downey's Results. *Psychological Reports*, 83(2), 608–610. <https://doi.org/10.2466/pr0.1998.83.2.608>
- Layne, C. (1999). Sex and the Barnum Effect: Rationality versus Helpfulness. *Psychological Reports*, 85(1), 187–188. <https://doi.org/10.2466/pr0.1999.85.1.187>
- Macdonald, D. J., & Standing, L. G. (2002). Does Self-Serving Bias Cancel the Barnum Effect? *Social Behavior & Personality: An International Journal*, 30(6), 625–630. <https://doi.org/10.2224/sbp.2002.30.6.625>
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52(1), 81–90. <https://doi.org/10.1037/0022-3514.52.1.81>
- McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of personality*, 60(2), 175–215. <https://doi.org/10.1111/j.1467-6494.1992.tb00970.x>
- Meehl, P. E. (1956). Wanted—a good cook-book. *American Psychologist*, 11(6), 263–272. <https://doi.org/10.1037/h0044164>
- Merrrens, M. R., & Richards, W. S. (1970). Acceptance of Generalized versus “Bona Fide” Personality Interpretation. *Psychological Reports*, 27(3), 691–694. <https://doi.org/10.2466/pr0.1970.27.3.691>
- Michels, P. J., & Layne, C. (1980). Inventory Responding Models People's Acceptance of Feedback “Derived” from Tests and from Interviews. *Journal of Personality Assessment*, 44(3), 302–306. https://doi.org/10.1207/s15327752jpa4403_14
- Mortaza Zare & Carol Flinchbaugh (2019) Voice, creativity, and big five personality traits: A meta-analysis. *Human Performance*, 32(1), 30–51. <https://doi.org/10.1080/08959285.2018.1550782>
- Nikolaou, I., & Foti, K. (2018). Personnel selection and personality. In V. Zeigler-Hill & T. K. Shackelford (Eds), *The SAGE handbook of personality and individual differences*, (vol. 3, pp. 458–474).
- Nikolaou, I., & Foti, K. (2018). Personnel selection and personality. In I. Nikolaou, & K. Foti *Personnel Selection and Personality* (pp. 458–474). SAGE Publications Ltd, <https://dx.doi.org/10.4135/9781526470300>
- Ozer, D. J., & Reise, S. P. (1994). Personality assessment. *Annual Review of Psychology*, 45, 357–388. <https://doi.org/10.1146/annurev.ps.45.020194.002041>
- Piper-Terry, M. L., & Downey, J. L. (1998). Sex, Gullibility, and the Barnum Effect. *Psychological Reports*, 82(2), 571–576. <https://doi.org/10.2466/pr0.1998.82.2.571>
- Rogers, P., & Soule, J. (2009). Cross-Cultural Differences in the Acceptance of Barnum Profiles Supposedly Derived From Western Versus Chinese Astrology. *Journal of Cross-Cultural Psychology*, 40(3), 381–399. <https://doi.org/10.1177/0022022109332843>
- Snyder, C. R., Shenkel, R. J., & Lowery, C. R. (1977). Acceptance of personality interpretations: The “Barnum effect” and beyond. *Journal of Consulting and Clinical Psychology*, 45, 104–114. <https://doi.org/10.1037/0022-006X.45.1.104>
- Snyder, C. R. (1974). Why horoscopes are true: The effects of specificity on acceptance of astrological interpretations. *Journal of Clinical Psychology*, 30(4), 577–580. [https://doi.org/10.1002/1097-4679\(197410\)30:4<577::AID-JCLP2270300434>3.0.CO;2-8](https://doi.org/10.1002/1097-4679(197410)30:4<577::AID-JCLP2270300434>3.0.CO;2-8)
- Stachnik, T., & Stachnik, B. (1980). Acceptance of Non-Specific Astrological Personality Descriptions: An Empirical Demonstration. *Psychological Reports*, 47(2), 537–538. <https://doi.org/10.2466/pr0.1980.47.2.537>
- Standing, L., & Keays, G. (1987). Do the Barnum Effect and Paranormal Belief Involve a General Gullibility Factor? *Psychological Reports*, 61(2), 435–438. <https://doi.org/10.2466/pr0.1987.61.2.435>
- Verberne, T. J. P. (1974). Forer Revisited: Acceptability of a Fake Personality Sketch. *Psychological Reports*, 34(3), 1001–1002. <https://doi.org/10.2466/pr0.1974.34.3.1001>
- Young, H. R., Glerum, D. R., Wang, W., & Joseph, D. L. (2018). Who are the most engaged at work? A meta-analysis of personality and employee engagement. *Journal of Organizational Behavior*, 39(10), 1330–1346. <https://doi.org/10.1002/job.2303>
- Zell, E., & Lesick, T. L. (2022). Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality*, 90(4), 559–573. <https://doi.org/10.1111/jopy.12683>
- Zell, E., Strickhouser, J. E., Sedikides, C., & Alicke, M. D. (2020). The better-than-average effect in comparative self-evaluation: A comprehensive review and meta-analysis. *Psychological Bulletin*, 146, 118–149. <https://doi.org/10.1037/bul0000218>