BOGDAN ZAWADZKI
University of Warsaw
Faculty of Psychology
JAN STRELAU
University of Social Sciences and Humanities
Faculty of Psychology

# ON THE (IN)VALIDITY
# OF THE GENERAL FACTOR OF PERSONALITY (GFP)

The paper presents research on the validity of the General Factor of Personality (GFP), extracted on the basis of NEO-FFI scores. Analysis was based on data obtained in four groups consisting, in total, of 1906 participants (878 women and 1028 men aged from 18 to 66 years). Scores on Five-Factor Model (FFM) scales and GFP scores were analyzed against 15 selected criterion variables: the intensity of PTSD symptoms, schizotypal personality symptoms, general self-efficacy, conflict management styles (integrating, dominating, avoiding, compromising, obliging), and behavioral risk factors of somatic diseases (CND and cancer): Type A, Type 1, Type 2, and Type 4 personality as well as depressiveness, hostility, and submissiveness. The results indicated the structural weakness of the GFP (low reliability, unacceptable fit of the one-factor model to the structure of NEO-FFI scales) and a high diversity of GFP validity. Comparative analysis of criterion variables for which both FFM and GFP scores had high validity with cases of variables showing high validity of FFM scores and low validity of the GFP allowed to identify the causes of the diversity of GFP validity. It is a result of low intercorrelations among FFM scales, which frequently creates differences between the pattern of intercorrelations of FFM scales and the pattern of their correlations with criterion variables. In the discussion the limited theoretical and empirical meaning of the GFP was underlined, especially when compared to the general $g$ factor, used in research on intelligence, extracted from highly intercorrelated ability tests or even personality scales, which are based on highly correlated lower level subdimensions.

**Keywords:** General Factor of Personality (GFP), Five-Factor Model of Personality (FFM).

## INTRODUCTION

According to the concept of the Five-Factor Model of Personality (FFM; see Costa & McCrae, 1992), the structure of personality can be comprehensively characterized by five factors: Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness, which fulfill the criteria of basic personality traits, namely: biological basis, reality, universality, invariance, and pervasiveness (Zawadzki, Strelau, Szczepaniak, & Śliwińska, 1998). Those traits were assumed to be orthogonal or intercorrelated to a very low extent, which implied that they cannot be aggregated into higher order factors. The first attempt to extract higher order factors was undertaken by Digman (1997), who found two factors of this kind in the structure of the FFM: **α (alpha)**, comprising Neuroticism (frequently recoded into Emotional Stability in subsequent studies due to negative factor loading), Conscientiousness, and Agreeableness, and **β (beta)**, comprising Extraversion and Openness. DeYoung, Peterson, and Higgins (2001) replicated the solution with two higher order factors, but labeled them, respectively, as **Stability** and **Plasticity**. These two studies initiated the search for higher order factors in the structure of personality traits and led to the identification of the **General Factor of Personality** (GFP). Musek (2007), who replicated the results obtained by Digman as well as DeYoung and colleagues, showed that the Stability and Plasticity factors were intercorelated, which made it possible to extract a more general personality factor, later named the *p* factor, by analogy with the *g* factor in the structure of mental ability tests. This idea was taken up by Rushton, who, in a series of papers, demonstrated the existence of the general factor in all available self-report personality measures (see, for example, Rushton & Irwing, 2009a, 2009b, or the summary of all these analyses in Just, 2011). The strategy of extracting the general factor was based on previous analyses and started from the identification of a few higher order factors in the structure of personality traits and finally – on the basis of their intercorrelations – extracting the GFP as the second-order factor. A theory supporting the GFP was also proposed, drawing on the theory of evolution and on sociobiology (Rushton, Bons, & Yoon-Mi, 2008). According to this theory, the GFP could be considered the basic personality dimension, reflecting the different strategies of human adaptation. One pole of the GFP is characterized by emotional stability, extraversion, openness, agreeableness and conscientiousness (the pattern of traits labeled as "nEOAC"), and the opposite pole by neuroticism, introversion, low openness, low agreeableness and low conscientiousness (the profile of traits called "Neoac"). It was also suggested that the GFP was related to the level of intelli-

gence. Its adaptive importance was underlined, for instance in occupational settings (Rushton et al., 2008). Some studies confirmed these expectations, but Just (2011) suggests that they need further replication, especially broader validation studies demonstrating the importance of the GFP for occupational success or its relationship with the socioeconomic status. The problem of GFP validation is also undertaken in this paper; however, we have not focused on the verification of Rushton's concept, but mainly on the methodological underpinnings of the GFP. The main aim was to demonstrate the validity of the GFP by relating it to several psychological constructs, assessed by self-report measures. This analysis was inspired by the critical discussion of the GFP and by previous studies on its validity, especially those advocating the importance of direct comparisons of predictions based on FFM scales and the GFP (see Van der Linden, te Nijenhuis, & Bakker, 2010).

It should be emphasized that the concept of the GFP was strongly criticized from the very beginning by researchers specializing in the psychology of personality and individual differences (see Holden & Marjanovic, 2012; Just, 2011). The main argument against it was the weakness of general factor measurement, essentially distinguishing personality inventories from tests of general intelligence, leading to the conclusion that the $p$ factor should be considered as a research artifact connected with self-esteem or social desirability (the self-presentation pattern). It was also underlined that the general factor saturates personality domains to different degrees and that it is mainly related to temperamental traits (Zawadzki & Strelau, 2010), which means it may be estimated on the basis of neuroticism and extraversion (Riemann & Kendler, 2010). This conclusion was also based on Musek's (2007) study, in which the GFP was extracted, also as a first order factor, directly from FFM traits, with the highest loadings found on Neuroticism, Extraversion, and Conscientiousness and much lower on Agreeableness and Openness. Thus, the results of these studies disconfirmed the general character of the GFP. Still, particularly important are the studies indicating low intercorrelations among FFM scales (essentially different from mental ability tests) and demonstrating the weakness of $p$ factor measurement. It was therefore suggested (Revelle & Wilt, 2013) that alternative structural models, such as the Spearmanian model (the general factor and specific dimensions) should be applied instead of the hierarchical one. It was also pointed out that such models confirmed the existence of the $g$ factor in the structure of mental abilities and at the same time disconfirmed the necessity of introducing the $p$ factor in the structure of FFM traits (Ashton, Lee, Goldberg, & de Vries, 2009). Ashton and colleagues (2009) claimed that the general factor is not neces-

sary to explain the correlations among FFM scales and suggested that these correlations are due to the same-sign blends of orthogonal traits (same-sign correlations between subdimensions and traits). According to this hypothesis, residual correlations among FFM scales stem from the insufficiently discriminative theoretical conceptualization of basic personality traits or inappropriate balance among subdimensions within scales. In this particular case it means that, for example, some of the subdimensions of Neuroticism should correlate negatively with Extraversion and others positively (and similarly for Extraversion). When appropriate balance of subdimensions is achieved, the total correlation between the two scales should be equal to zero, with no need for a higher order factor. In this sense, the correlations among FFM scales are in fact residual ones, and their low values do not point to the existence of any substantial personality dimension.

The assumption behind the analysis presented in this paper was the thesis that the magnitude of the correlations among FFM scales is crucial for the internal as well as external validity of the GFP. However, the present study differs from previous ones with regard to three aspects. First, what is essential for internal validity is the extent to which FFM scales are valid and proportionate measures of the GFP; otherwise the loadings on the $p$ factor will be low for all scales or substantial only for some of them. In the latter case, the GFP will exhibit measurement weakness or inappropriate content structure – and so it will have a specific rather than general character. This requires an analysis of the GFP as the first order factor in the structure of FFM scales, with no need for hierarchical analysis via intermediate order factors. Second, low intercorrelations among FFM scales may result in the variability of GFP validity when it is related to different criterion variables. In other words, the general factor may demonstrate high or low validity depending on the pattern of correlations between FFM scales and criterion variables. It should be underlined that for mental ability tests high validity of the $g$ factor is usually found even though validity differs between component scales (however, due to the high intercorrelations among them, finding both positive and negative correlations between some of them and criterion variables is rare and not very probable; see for example Brzeziński & Hornowska, 1993; Matczak, Jaworowska, Ciechanowicz, & Stańczak, 2006). Third, in order to test the validity of the GFP, we took a broad set of different psychological variables into account: symptoms of PTSD and schizotypal personality, general self-efficacy, styles of conflict management, and behavioral potential risk factors for somatic diseases. All of them are related to FFM traits, but these relationships differ in regard to magnitude and direction, which enables a more systematic analysis of GFP validity.

Specific hypotheses concerning the relationships between FFM traits and criterion variables are presented elsewhere (see Bożko, 2000; Osuch, 2005; Parnowska, 2011), so in this paper we only formulated expectations regarding:

– the weakness of the measurement of the general factor and its saturation by only some personality domains;

– the variability of GFP validity depending on the specificity of criterion variable, which suggests the limited empirical diagnostic value of the GFP.

## METHOD

The results obtained in previous studies within different projects were reanalyzed. In all of the studies, FFM traits were assessed using the NEO-FFI (Costa & McCrae, 1992; Zawadzki et al., 1998). The GFP was calculated as a simple sum of standardized scores on FFM scales, as in intelligence tests (after recoding the scores on the Neuroticism scale into Emotional Stability; see Rushton et al., 2008), as well as through exploratory factor analysis (using the Principal Axes method) and by saving the factor scores via the regression method (see Zawadzki & Strelau, 2010). In the validity analysis, FFM scores as well as GFP scores were related to other psychological variables, such as symptoms of PTSD and schizotypal personality, general self-efficacy, styles of conflict management, and behavioral potential risk factors for somatic diseases.[1]

### Samples and Instruments

The first analysis was focused on the validity of FFM scales and GFP indicators with regard to the intensity of PTSD symptoms and general self-efficacy (see Kaczmarek & Zawadzki, 2012). In this study, the intensity of PTSD symptoms was assessed using the PTSD-F inventory (Strelau, Zawadzki, Oniszczenko, & Sobolewski, 2002) and general self-efficacy was measured using the Polish version of the General Self-Efficacy Scale (GSES; Juczyński, 2009; Luszczynska, Scholz, & Schwarzer, 2005). The demographic characteristics of this group (labeled Sample 1) as well as of other samples are presented in Table 1.

---

[1] The notion of criterion variables was applied in this paper only to mean the reference variables that were used to validate the GFP. All the analyses presented were focused on construct validity and should not be considered as classic criterion-oriented validation.

Table 1
*Demographic Characteristics of the Samples*

| Sample | $N$ | Gender | Age: range | Age: $M$ ($SD$) |
|---|---|---|---|---|
| 1 | 1132 | 464 F / 668 M | 18-66 | 36.00 (13.57) |
| 2 | 402 | 214 F / 188 M | 18-50 | 22.42 (4.70) |
| 3 | 172 | 101 F / 71 M | 19-40 | 21.26 (2.43) |
| 4 | 200 | 99 F / 101 M | 20-25 | 22.64 (1.46) |
| Samples 1-4 combined | 1906 | 878 F / 1028 M | 18-66 | 31.43 (13.12) |

The symptoms of schizotypal personality were the reference variable in the second analysis. Data were obtained by Parnowska (2011), who applied the Polish version of the Schizotypal Personality Questionnaire (SPQ) developed by Raine (1991). The third analysis was based on data obtained by means of the Rahim Organizational Conflict Inventory (ROCI-II; Rahim, 1983) and was focused on the relationships between the FFM and the GFP and conflict management styles (integrating, dominating, compromising, avoiding, and obliging). Polish version of ROCI-II was developed by Osuch (2005), who also conducted the studies in Sample 3. The last analysis was done on data obtained by Bożko (2000) in Sample 4. In this analysis, the relationships between FFM traits and behavioral potential risk factors for somatic diseases (CHD and cancer) were examined. Several inventories were applied: Type A Behavior Inventory assessing the symptoms of Type A behavior (Wrześniewski, 1990), the Polish version of the Short Interpersonal Reaction Inventory (Grossarth-Maticek & Eysenck, 1990) diagnosing the symptoms of Type 1 personality (cancer-prone), Type 2 personality (CHD-prone), and Type 4 personality (healthy personality – in the analysis, scale 4B was applied due to its reliability much higher than that of scale 4A), and the Inventory of Personality Patterns (see Zawadzki & Radzikowska, 2006), assessing hostility, depressiveness, and submissiveness, understood as potential personality risk factors for somatic diseases – CHD and lung cancer. Altogether, the combined sample (1-4) consisted of 1906 participants (878 women and 1028 men) aged 18-66 years.

## The Procedure
## of GFP Validity Analysis

The main procedure applied was based on a comparison of correlations with criterion variables obtained for FFM scales and GFP indicators. The formal test

of incremental validity for all FFM scales and the GFP via linear regression was impossible due to the multicollinearity of the GFP indicator and a majority of FFM scales. Comparisons of the multiple correlations of FFM scales and simple correlation of the GFP with criterion variables showed an advantage of the FFM scales due to the appropriate fit of regression weights for the predicted variables. For this reason, a simplified analysis was carried out, based on a comparison between the correlation with a given criterion variable obtained for one FFM scale demonstrating the highest validity (printed in bold in Table 3) and the corresponding correlation obtained for the GFP. The difference between the two correlations was tested using a two-tailed *t*-test (see Cohen & Cohen, 1983). The formal analysis of incremental validity, even for one selected FFM scale and the GFP, was also impossible or misleading due to: (1) the high multicollinearity of both independent variables, especially evident for Neuroticism, (2) the significant increment of explained variance for both the GFP and the FFM scale demonstrating the highest validity (in the case of significant relationships between criterion variables and all or most FFM traits), and (3) significant semipartial correlations of the GFP, accompanied by its low simple correlations with criterion variables or even by a change of the sign of semipartial correlations in comparison to simple correlations, especially in the case of Agreeableness. This might suggest an effect of capitalization on chance in a particular sample, which could be a difficult matter for theoretical interpretation. For these reasons, the comparative analysis was restricted only to comparisons of simple correlations of one selected FFM scale and the GFP with criterion variables.

## RESULTS

### The Internal Validity of the GFP

Analysis began with extracting the General Factor of Personality from data obtained in each sample and all groups combined into an overall sample. Table 2 presents factor loadings of NEO-FFI scales for the general factor, eigenvalues, medians of correlations among FFM scales, and GFP reliabilities: the values of Cronbach's α coefficient calculated for five scales.

Table 2

*The Factor Loadings and β Weights of FFM Scales (From the NEO-FFI) for the GFP*

| Factor loadings of FFM scales | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Samples 1-4 combined |
|---|---|---|---|---|---|
| N | -.68 | -.61 | -.69 | -.68 | -.65 |
| E | .59 | .69 | .67 | .68 | .63 |
| O | .08 | .28 | .30 | .11 | .14 |
| A | .37 | .39 | .33 | .30 | .36 |
| C | .54 | .56 | .40 | .63 | .52 |
| Eigenvalue | 1.23 (24.6%) | 1.39 (27.8%) | 1.29 (25.7%) | 1.42 (28.5%) | 1.23 (24.6%) |
| Correlations among FFM scales (range) | *Mdn* = .19 (-.03 to .40) | *Mdn* = .25 (.06 to .42) | *Mdn* = .20 (.11 to .50) | *Mdn* = .20 (-.02 to .46) | *Mdn* = .20 (-.03 to .41) |
| Cronbach's α | .54 | .62 | .59 | .59 | .56 |
| β weights of FFM scales | Sample 1 | Sample 2 | Sample 3 | Sample 4 | Samples 1-4 combined |
| N | -.52 | -.36 | -.50 | -.43 | -.47 |
| E | .38 | .50 | .47 | .43 | .44 |
| O | .04 | .10 | .13 | .04 | .06 |
| A | .18 | .17 | .15 | .12 | .18 |
| C | .32 | .30 | .19 | .35 | .30 |

*Note.* N – Neuroticism, E – Extraversion, O – Openness, A – Agreeableness, C – Conscientiousness. N scores were recoded into Emotional Stability for the calculation of Cronbach's α and the simple sum of FFM scales (GFP). Differences in the reliability coefficient among samples were not statistically significant ($\chi^2 = 4.90$ for $df = 3$, just like those between gender subgroups within the combined sample 1-4, $\chi^2 = 0.11$ for $df = 1$, and the three age subgroups, $\chi^2 = 0.11$ for $df = 2$; see Hakstian & Whalen, 1976). The reliability coefficients of NEO-FFI scales in the combined sample (1-4) were: .84 (N), .78 (E), .69 (O), .68 (A), & .82 (C). The general factor (GFP) was extracted using the Principal Axes Method for standardized scores of FFM scales. For the simple sum of scales, β weights were equal to .34 (Sample 1), .32 (Sample 2), and .33 (Sample 3, 4, and combined samples) for all FFM scales. Tucker's coefficients of factor congruence between particular samples and the combined sample: from .982 to .997 (in direct sample comparisons the lowest coefficient was found for Samples 1 and 3: .970).

The results are in line with expectations. First, data showed that the general factor is saturated only by some personality traits – mainly by Neuroticism and Extraversion, and – to lower extent – by Conscientiousness, Agreeableness and especially by Openness. The high value of factor similarity coefficients indicates the invariance of the factor structure across all groups. The results are fully consistent with the data obtained in our previous studies, showing that the GFP comprises only temperamental traits and cannot be attributed to gender or age

differences because very similar findings were obtained for data corrected for these demographic variables (Zawadzki & Strelau, 2010). Second, the reliability coefficients (with a value around .60) as well as the eigenvalues (about 25% of explained variance for FFM scales) point to the weakness of measurement of the general factor. The main cause of this is low intercorrelations among NEO-FFI scales (median equal to .20). As a result, an attempt to explain the variance of FFM traits by means of one common latent dimension did not find sufficient empirical support – the model with one first order factor in confirmatory factor analysis did not provide acceptable fit (in combined samples 1-4): $\chi^2 = 86.40$, $df = 5$, $p = .00$, RMSEA = .091, GFI = .993. In all subsequent analyses – in spite of the evident weakness of the general factor – two indicators of the GFP were calculated: factorial (obtained in exploratory factor analysis) and simple sum of standardized scores of NEO-FFI scales. The latter was recommended by Rushton and colleagues (2008), similarly to the general intelligence score in mental ability tests, but in this analysis it was considered only as a supplementary one. For both indicators of the GFP, β weights were calculated showing the contribution of each FFM scale to the GFP by linear regression with NEO-FFI scales as predictors of the general factor.

## *Validity of the GFP*
## *– Analysis of Correlations With Criterion Variables*

Table 3 presents the correlations of FFM scales and both indicators of the GFP (the sum of scores and the factorial score) with all criterion variables. The FFM scales that significantly predicted reference variables in regression analysis (which did not take the GFP into account) are indicated as well as the results of comparisons of criterion correlations obtained for the factorial GFP score and one FFM scale demonstrating the highest validity.

Only three (out of 15) correlations of factorial GFP score with criterion variables were not statistically significant and their median was equal to .34 (with high similarity of correlation obtained for simple sum and factorial GFP score). This result, however, only partly confirms the validity of the GFP. Full confirmation requires analysis taking into account the validity of FFM scales and careful analysis of criterion variables demonstrating high and low validity of the GFP. The same analysis also showed the high dispersion of correlation coefficients for the GFP, ranging from .06 to .70, which is in line with our theoretical expectations.

Table 3

*The Validity of FFM Scales and the GFP – Correlations With Symptoms of PTSD, Schizotypal Personality, Conflict Management Style,s and Behavioral Potential Risk Factors for Somatic Diseases (CHD and Cancer)*

| FFM scales | Sample 1: PTSD | Sample 1: Self-efficacy | Sample 2: Schizo-typal personality | Sample 3: Integrating style | Sample 3: Avoiding style |
|---|---|---|---|---|---|
| N | **.44*** | **-.49*** | **.48*** | -.34* | **.31*** |
| E | -.18* | .39* | -.35* | .34* | -.22* |
| O | .01 | .07* | .01 | .33* | -.24* |
| A | -.05* | .05 | -.30* | **.39*** | .05 |
| C | -.12* | .36* | -.30* | .27* | .03 |
| GFP | **-.34*** | **.53*** | **-.49*** | **.48*** | **-.27*** |
| Sum of scales | -.26* | .46* | -.45* | .54* | -.22* |
| FFM predictors | N | nEoAC | NeOa | nEOA | No |
| GFP vs. FFM | FFM (.01) | GFP (.01) | *ns* | *ns* | *ns* |

| FFM scales | Sample 3: Dominating style | Sample 3: Obliging style | Sample 3: Com-promising style | Sample 4: Type 1 | Sample 4: Type 2 |
|---|---|---|---|---|---|
| N | -.08 | **.18*** | -.03 | **.37*** | **.38*** |
| E | .16* | -.12 | .17* | -.33* | -.22* |
| O | .05 | -.16* | .10 | -.23* | -.06 |
| A | **-.44*** | -.11 | **.29*** | -.13 | -.16* |
| C | .03 | -.12 | .06 | -.18* | -.09 |
| GFP | **.06** | **-.17*** | **.16*** | **-.36*** | **-.31*** |
| Sum of scales | -.04 | -.15* | .21* | -.32* | -.30* |
| FFM predictors | eA | NoA | A | NeoA | N |
| GFP vs FFM | FFM (.01) | *ns* | *ns* | *ns* | *ns* |

| FFM scales | Sample 4: Type 4 | Sample 4: Type A | Sample 4: Depressiveness | Sample 4: Hostility | Sample 4: Submissiveness |
|---|---|---|---|---|---|
| N | **-.55*** | .07 | **.75*** | .32* | .13* |
| E | .41* | .33* | -.56* | .03 | -.12* |
| O | .13 | .15* | -.01 | .06 | -.23* |
| A | .22* | **-.40*** | -.08 | **-.45*** | **.32*** |
| C | .31* | .08 | -.38* | -.14* | .02 |
| GFP | **.55*** | **.10** | **-.70*** | **-.22*** | **-.07** |
| Sum of scales | .53* | .03 | -.58* | -.27* | -.05 |
| FFM predictors | nE | NEa | Ne | NEa | NoA |
| GFP vs. FFM | *ns* | FFM (.01) | *ns* | FFM (.01) | FFM (.01) |

*Note*. FFM predictors – NEO-FFI scales showing significant semipartial correlations in regression analysis with criterion variables (analysis without the GFP). Significant correlation coefficients (at $p < .05$) are marked with an asterisk. The highest correlations of FFM scales as well as the correlations of factorial GFP indicator are printed in bold. To compare the correlations of the GFP and a selected FFM scale with criterion variables, a two-tailed *t*-test was applied (see Cohen & Cohen, 1983).

As we suggested in Introduction, the GFP should demonstrate high validity when (1) FFM scales also show high correlation with criterion variables and (2) the magnitude and direction of these correlations are in line with the factor loadings or β weights obtained for the GFP. Thus, it should be consistent with the pattern of FFM traits coded as "nEOAC" or with its opposite: "Neoac." This is exemplified by high correlations for FFM scales and the GFP with integrating style of conflict management ("nEOA" pattern in regression analysis of FFM scales), general self-efficacy ("nE**o**AC"), symptoms of schizotypal personality ("Ne**O**a"**)**, Type 1 ("nE**o**A"), and Type 4 of personality ("nE") or depresiveness ("Ne"; scales with correlations not congruent with the pattern of traits describing both poles of the GFP are marked with bold type). The reverse correlation found for Agreeableness or Openness does not decrease the validity of the GFP provided that the correlations of other scales are sufficiently high and consistent with the pattern of traits for one pole of the GFP. The reason is that the contribution of these scales to the GFP (especially Openness) is relatively low (see β weights in Table 2). It should also be noted, that only in one case (general self-efficacy) was the correlation of GFP significantly higher than the highest correlation obtained for FFM scales, but for other criterion variables the validity of the GFP was essentially identical to the validity of the selected FFM scale.

All other criterion correlations demonstrate rather low validity of the GFP. Relatively low correlations for the GFP were obtained in three different sorts of cases. The first one refers to low validity of FFM scales, which was found for the avoiding and obliging styles of conflict management. The reason for the low validity of the GFP is obvious – the lack of validity of FFM scales causes the invalidity of aggregate scores. The second case refers to the low validity of the GFP caused by the differential validity of FFM scales, as when only some of them show sufficient validity (from the theoretical point of view, when a criterion variable is related to only one FFM trait). It was found even for FFM scales with a high contribution to the GFP (high β weight), as in the case of the intensity of PTSD symptoms or Type 2 personality ("N"), but was especially evident when the contribution of a particular FFM scale was low (low β weight), as for the compromising style of conflict management (only "A"). The low validity of the GFP stems from the attenuation of the validity of a particular FFM scale when it is aggregated into the general score together with scales that have insufficient validity. The lowest correlations of the GFP were found in the last case, when the pattern of criterion correlations of FFM scales was different from the "nEOAC" or "Neoac" patterns. These findings were obtained for the dominating style of conflict management ("**Ea**"), Type A ("N**E**a"), hostility ("N**E**a"),

and submissiveness ("No**A**"). For all these criteria substantially high correlations with FFM scales were found; they were significantly higher than the correlations found for the GFP. This part of analysis documented the low validity of the GFP, despite the high validity of FFM scales. Thus, not only is the criterion correlation of FFM scales attenuated, but the opposite correlations of FFM scales aggregated into the general score seem to completely erase the validity of the GFP. This happens if the pattern of correlations of FFM scales with a given criterion variable is opposite to the pattern of aggregation into the GFP, for example when the criterion variable is predicted by high neuroticism and high extraversion. After the aggregation of FFM scales, their contribution to the validity of the GFP is neutralized, depriving the general factor of empirical significance in spite of the high validity of source scales. The low correlations among FFM scales suggest that this may happen relatively frequently in personality studies, which means that for many criteria data might be obtained indicating high validity of FFM scales combined with low or close-to-zero validity of the GFP score.

## DISCUSSION

The results of analyses confirmed the theoretical expectations. First of all, the study showed the high dispersion of correlations of the GFP with criterion variables, which by itself may be interpreted against the validity of the GFP. All analyses make it possible to explain the causes of the dispersion of validity coefficients, that is, to indicate what requirements were fulfilled when the general factor reached high validity and under what conditions the GFP indicator lost its validity. The main causes concern the rather obvious requirement of the validity of component scales and the consistency of the magnitude and direction of criterion correlations with the β weights of source scales obtained when extracting the GFP. Especially important, however, are the data indicating a high validity of component FFM scales accompanied by a low validity of the general factor. In such cases, the aggregation of FFM scales scores, based on their internal structure and different from the pattern of correlations with criterion variables, seemed to attenuate the validity of component scales and decrease the validity of the GFP. The cause of the validity as well as invalidity of the general factor lies in the very operation of aggregation, if it is not congruent with the prediction of criterion variables. This effect does not exist in regression analysis, which aggregates the scores of scales on the basis of their empirical importance and the direction of relationships between independent and criterion variables. Does this

mean that aggregation should be abandoned in personality studies? After all, in every case aggregation leads to a loss of information about the specificity of behavior. However, in the psychology of individual differences aggregation is essential and common: items are aggregated into scales and subdimensions into a general trait – for example, the scores of mental ability tests are aggregated into the general intelligence indicator and subdimensions such as depression or anxiety into a broad personality trait such as neuroticism. Which aggregation is reasonable or – in other words, on which level should it stop? The criterion is relatively simple and requires appropriately high intercorrelations of component scales and refraining from further aggregation when the correlations among the higher order variables are relatively low or close to zero (Nowakowska, 1975). In the case of very high correlation it could be argued that a new variable is being assessed, not reducible to component variables, and the validity of the general factor should be high for various criterion variables. Cases of the validity of component variables being the reverse of the pattern of aggregation – which decreases the validity of the general factor score – should also be very rare when the scales are highly intercorrelated. The idea of the GFP obviously violates this simple rule – correlations among FFM scales are too low to generate a general factor and forced aggregation leads to a considerable dispersion of the validity coefficient of general scores. Data indicating high validity of the GFP are not really conclusive. The arbitrary sum of orthogonal scales would also demonstrate high validity if the component scales were highly valid. However, it would immediately lose validity if a different criterion variable was considered. For all such variables, different aggregates of source scales could be created, highly valid only for predicting that particular dependent variable. From a more general perspective, accepting this paradigm would mean that we have to remain only on the level of source variables, being aggregated into appropriate general scores depending on the aims of the study – and if we were to remain at the level of component scores, the sense of their arbitrary aggregation would be very doubtful.

The crucial problem for the *p* factor is summarized in the question: Is the magnitude of intercorrelations among FFM scales sufficient for aggregation to make sense or – from the theoretical perspective – does the GFP reflect an important personality dimension (assuming that FFM traits comprehensively characterize the structure of personality)? First of all, it should be stressed that inter-correlation among FFM scales really exists. It was nicely summarized by Szarota (1995, p. 248), who commented on the findings of lexical studies on personality dimensions as follows: "significant correlations between the Big Five dimen-

sions have been reported so frequently that even theoreticians have ceased to believe in their orthogonality." On the other hand, it is suggested that correlations among FFM scales are too low to indicate that a substantial psychological variable is being measured, especially in comparison to mental ability tests (Ashton et al., 2009; Revelle & Wilt, 2013). Similar findings were obtained in the present study, which indicates the weakness of the GFP. It is also suggested that correlations among FFM scales may be considered as only residual correlations, stemming from the unidirectional relationships between subdimensions and general traits (this mainly refers to Neuroticism and Extraversion; see Riemann & Kandler, 2010), which means that there is no evidence for the existence of important higher order psychological variables (Ashton et al., 2009). The results of studies carried out by Musek (2007) or Riemann and Kandler (2010) as well as our findings presented elsewhere (Zawadzki & Strelau, 2010) show a heterogeneous saturation of FFM traits with the general factor. Identical data were obtained in this study, indicating that the general factor is in fact not a general but rather a specific one and comprises mainly temperamental traits (Neuroticism and Extraversion). This is a very important argument for the claim that temperament is a pivotal factor in the structure of personality – and, we believe, the theoretical significance of GFP analyses in fact ends at this claim. The procedure of aggregating variables makes very limited empirical sense when it is not based on solid theoretical grounds and does not fulfill strict methodological requirements.

## REFERENCES

Ashton, M. C., Lee, K., Goldberg, L. R., & de Vries, R. E. (2009). Higher-order factors of personality: Do they exist? *Personality and Social Psychology Review, 13*, 79-91.

Bożko, B. (2000). *Depresyjność i wybrane składniki wzoru zachowania Typu A i Typu C a struktura podstawowych cech osobowości* (Unpublished master's thesis). The Faculty of Psychology of the University of Warsaw.

Brzeziński, J., & Hornowska, E. (Eds.). (1993) *Skala inteligencji Wechslera. WAIS-R. Polska adaptacja. standaryzacja. normalizacja i wykorzystanie w diagnostyce psychologicznej*. Warsaw: Wydawnictwo Naukowe PWN.

Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences.* Hillsdale. NJ: Lawrence Erlbaum Associates.

Costa, P. T., Jr., & McCrae, R. R. (1992*). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa. FL: Psychological Assessment Resources.

DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2001). Higher-order factors of the Big Five predict conformity: Are there neuroses of health? *Personality and Individual Differences, 33*, 533-552.

Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology, 73*, 1246-1256.

Grossarth-Maticek, R., & Eysenck, H. J. (1990). Personality, stress and disease: Description and validation of a new inventory. *Psychological Reports, 66*, 355-373.

Hakstian, A. R., & Whalen, T. E. (1976). A *k*-sample significance test for independent alpha coefficients. *Psychometrika, 41*, 219-231.

Holden, R. R., & Marjanovic, Z. (2012). A putatively general factor of personality (GFP) is not so general: A demonstration with the NEO PI-R. *Personality and Individual Differences, 52*, 37-40.

Juczyński, Z. (2009). *Narzędzia pomiaru w promocji i psychologii zdrowia*. Warsaw: Pracownia Testów Psychologicznych PTP.

Just, C. (2011). A review of the literature on general factor of personality. *Personality and Individual Differences, 50*, 765-771.

Kaczmarek, M., & Zawadzki, B. (2012). Exposure to trauma, emotional reactivity and its interaction as predictors of the intensity of PTSD symptoms in the aftermath of motor vehicle accidents (MVA). *Journal of Russian and East European Psychology, 50*, 47-64.

Luszczynska, A., Scholz. U., & Schwarzer, R. (2005). The general self-efficacy scale: Multicultural validation studies. *The Journal of Psychology, 139*, 439-457.

Matczak, A., Jaworowska, A., Ciechanowicz, A., & Stańczak. J. (2006). *Bateria Testów APIS-Z*. Warsaw: Pracownia Testów Psychologicznych PTP.

Musek, J. (2007). A general factor of personality: Evidence for the Big One in the Five-Factor Model. *Journal of Research in Personality, 41*, 1213-1233.

Nowakowska, M. (1975). *Psychologia ilościowa z elementami naukometrii*. Warsaw: Wydawnictwo Naukowe PWN.

Osuch, K. (2005). *Style zachowania w konflikcie a cechy osobowości* (Unpublished master's thesis). The Faculty of Psychology of the University of Warsaw.

Parnowska, D. (2011). *Genetyczne i środowiskowe uwarunkowania struktury schizotypii u osób zdrowych* (Unpublished doctoral dissertation). The Institute of Psychiatry and Neurology, Warsaw.

Rahim, M. A. (1983). A measure of styles of handling interpersonal conflict. *Academy of Management Journal, 26*, 368-376.

Raine, A. (1991). The SPQ: Scale for the assessment of schizotypal personality based on DSM-III-R criteria. *Schizophrenia Bulletin, 17*, 555-564.

Revelle, W., & Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality, 47*, 493-504.

Riemann, R., & Kandler, C. (2010). Construct validation using multitrait-multimethod-twin data: The case of a general factor of personality. *European Journal of Personality, 24*, 258-277.

Rushton, J. P., Bons, T. A., & Yoon-Mi, H. (2008). The genetics and evolution of the general factor of personality. *Journal of Research in Personality, 42*, 1173-1185.

Rushton, J. P., & Irwing, P. (2009a). A General Factor of Personality in 16 sets of the Big Five, the Guilford-Zimmerman Temperament Survey, the California Psychological Inventory, and the Temperament and Character Inventory. *Personality and Individual Differences, 47*, 558-564.

Rushton, J. P., & Irwing, P. (2009b). A General Factor of Personality (GFP) from the Multidimensional Personality Questionnaire. *Personality and Individual Differences, 47*, 571-576.

Strelau, J., Zawadzki, B., Oniszczenko, W., & Sobolewski, A. (2002). Kwestionariusz PTSD – wersja czynnikowa (PTSD-C). Konstrukcja narzędzia do diagnozy głównych wymiarów zespołu stresu pourazowego. *Przegląd Psychologiczny, 45*(2), 149-176.

Szarota, P. (1995). Polska Lista Przymiotnikowa (PLP): Narzędzie do diagnozy Pięciu Wielkich czynników osobowości. *Studia Psychologiczne, 33*, 227-256.

Van der Linden, D., te Nijenhuis, J., & Bakker, A. B. (2010). The General Factor of Personality: A meta-analysis of Big Five intercorrelations and a criterion-related validity study. *Journal of Research in Personality, 44*, 315-327.

Wrześniewski, K. (1990). Badanie Wzoru zachowania A przy użyciu polskiego kwestionariusza. *Przegląd Lekarski, 47*, 538-542.

Zawadzki, B., & Radzikowska, E. (2006). Próba kwestionariuszowej diagnozy potencjalnych osobowościowych czynników ryzyka raka płuca i choroby wieńcowej. *Psychologia – Etologia – Genetyka, 13*, 135-154.

Zawadzki, B., & Strelau, J. (2010). Structure of personality: Search for a general factor viewed from a temperament perspective. *Personality and Individual Differences, 49*, 77-82.

Zawadzki, B., Strelau, J., Szczepaniak, P., & Śliwińska, M. (1998). *Inwentarz Osobowości NEO-FFI Costy i McCrae – adaptacja polska. Podręcznik*. Warsaw: Pracownia Testów Psychologicznych PTP.