

BOGDAN ZAWADZKI

Uniwersytet Warszawski

Wydział Psychologii

JAN STRELAU

Szkola Wyzsza Psychologii Spolecznej

Wydział Psychologii

O (NIE)TRAFNOŚCI OGÓLNEGO CZYNNIKA OSOBOWOŚCI (GFP)

W pracy podjęto problem trafności Ogólnego Czynnika Osobowości (*General Factor of Personality*, GFP). Analiza bazowała na danych zebranych ogółem w grupie 1906 osób (878 kobiet i 1028 mężczyzn w wieku 18-66 lat). GFP wyodrębniono na podstawie wyników badania kwestionariuszem NEO-FFI, zaś wyniki skal badających cechy Pięciodzennikowego Modelu Osobowości (PMO) oraz czynnika GFP odniesiono do 15 wybranych zmiennych kryterialnych: nasilenia objawów PTSD, uogólnionego poczucia własnej skuteczności, objawów osobowości schizotypowej, stylów zachowania w konflikcie (konsens, unikanie, dominacja, uleganie i kompromis) oraz behawioralnych czynników ryzyka chorób somatycznych (choroby wieńcowej i nowotworowej): Typu A, Typu 1, Typu 2 i Typu 4 osobowości oraz depresyjności, uległości i wrogości. Uzyskane wyniki wskazały na słabość pomiaru GFP (niska rzetelność, nieadekwatne dopasowanie modelu jednoczynnikowego do struktury skal NEO-FFI) oraz dużą zmienność trafności GFP. Analiza porównawcza zmiennych kryterialnych, dla których zarówno GFP, jak i skale PMO charakteryzowała wysoka trafność z przypadkami zmiennych wykazujących niską trafność czynnika GFP przy wysokiej trafności skal źródłowych, pozwoliła na zidentyfikowanie przyczyn zróżnicowania trafności GFP. Wynika ona z niskiego skorelowania skal PMO, skutkującego odmiennością wzorca interkorelacji skal od wzorca ich skorelowania ze zmiennymi kryterialnymi. W dyskusji wskazano na ograniczony sens teoretyczny i empiryczny czynnika GFP, zasadniczo odmienny od stosowanego w badaniach nad inteligencją czynnika ogólnego *g*, generowanego z testów zdolności czy nawet skal osobowości, które są agregatami wysokoskorelowanych cech niższego rzędu.

Słowa kluczowe: Ogólny Czynniki Osobowości (GFP), Pięciodzennikowy Model Osobowości (FFM).

Adres do korespondencji: BOGDAN ZAWADZKI – Uniwersytet Warszawski, Wydział Psychologii, ul. Stawki 5/7, 00-183 Warszawa; e-mail: bogdan@psych.uw.edu.pl

Praca przygotowana w ramach Grantu Narodowego Centrum Nauki 2012/06/A/HS6/00340 „PTSD: Diagnoza Terapia Profilaktyka”.

WPROWADZENIE

Jedno z podstawowych twierdzeń Pięcioczynnikowego Modelu Osobowości (PMO; zob. Costa i McCrae, 1992) głosi, że strukturę osobowości wyczerpująco charakteryzuje pięć czynników: neurotyczność, ekstrawersja, otwartość na doświadczenie, ugodowość i sumienność, spełniających kryteria wymagane od podstawowych cech osobowości, a mianowicie biologiczności, uniwersalności, realności, niezmienniczości i ogólności (Zawadzki, Strelau, Szczepaniak i Śliwińska, 1998). Cechy te – z racji ich zakładanej ortogonalności, a faktycznie niskiego skorelowania – traktowano jako dalej już nieredukowalne, a więc przyjmowano tezę o nieistnieniu czynników wyższego rzędu. Pierwszy wyłom w tym sposobie myślenia stanowiły analizy Digmana (1997), który w strukturze cech PMO wyodrębnił dwa czynniki wyższego rzędu: α (alfa), obejmujący neurotyczność (z racji ujemnego ładunku często w pracach późniejszych rekodowaną do stabilności emocjonalnej), sumienność i ugodowość, oraz β (beta), w którego skład wchodzi otwartość i ekstrawersja. Analizy te zostały później zreplicowane przez DeYounga, Petersona i Higginsa (2001), którzy zinterpretowali treści obu czynników jako **stabilność** i **plastyczność**. Prace te zapoczątkowały poszukiwania czynników wyższego rzędu w strukturze cech osobowości, które doprowadziły do wyodrębnienia **Ogólnego Czynnika Osobowości** (*General Factor of Personality*, GFP). Musek (2007), który kontynuował te analizy, replikując wyniki Digmana oraz DeYounga i współpracowników, wykazał, że czynniki stabilności i plastyczności, z racji skorelowania, pozwalają na wyodrębnienie najbardziej ogólnej zmiennej osobowościowej, nazwanej później – na wzór czynnika *g* w strukturze zdolności umysłowych – czynnikiem *p*. Idea ta została podjęta przez Rushtona, który w kolejnych pracach wykazywał obecność czynnika ogólnego we wszystkich dostępnych narzędziach samoopisowych do badania osobowości (zob. przykładowo Rushton i Irving, 2009a, 2009b oraz podsumowanie tych analiz w pracy Just, 2011). Przyjęta strategia była wzorowana na tych pierwotnych analizach i polegała na wyodrębnianiu w strukturze cech kilku czynników wyższego rzędu oraz – na podstawie ich skorelowania – generowania ostatecznie GFP (jako czynnika drugiego rzędu). Przedstawiono także koncepcję uzasadniającą GFP, odwołującą się do teorii ewolucji i socjobiologii (Rushton i in., 2008). Zgodnie z tą koncepcją GFP jest podstawowym wymiarem osobowości odzwierciedlającym odmienne strategie adaptacyjne, których jeden biegun jest charakteryzowany przez zrównoważenie emocjonalne, ekstrawersję, otwartość, ugodowość i sumienność (wzór, który schematycznie może być opisany przez profil cech: „nEOAC”), zaś drugi poprzez neurotyczność, introwersję, niską otwartość, niską ugodowość oraz niską sumienność (profil cech: „Neoac”).

Sugerowano także powiązanie GFP z inteligencją oraz jego znaczenie przystosowawcze, np. w aktywności zawodowej (Rushton i in., 2008). Niektóre z tych analiz istotnie potwierdziły te hipotezy badawcze, ale jak uważa Just (2011) – wymagają one replikacji, w tym zwłaszcza szerszych analiz walidacyjnych, np. wykazania powiązania GFP ze statusem społecznym czy powodzeniem zawodowym. Problem walidacji GFP został zatem podjęty w tej pracy, ale nie z perspektywy weryfikacji tej koncepcji Rushtona, ale o bardziej metodologicznym charakterze. Miała ona na celu zilustrowanie wartości diagnostycznej GFP przy odniesieniu do różnorodnych zmiennych psychologicznych, badanych zasadniczo przy zastosowaniu metod kwestionariuszowych. Analizę podjęto w kontekście krytycznej argumentacji przeciwko GFP, w tym zwłaszcza uwzględniającej porównanie predykcji na podstawie poszczególnych cech oraz indeksu ogólnego (zob. Van der Linden, te Nijenhuis i Bakker, 2010).

Od samego początku bowiem koncepcja GFP spotkała się z nasiloną krytyką w psychologii różnic indywidualnych (zob. Holden i Marjanovic, 2012; Just, 2011). Głównym argumentem podnoszonym przez przeciwników jest słabość pomiaru czynnika ogólnego, zasadniczo różniąc kwestionariusze osobowości od testów od inteligencji ogólnej, wskutek czego indeks *p* jest traktowany jako artefakt badawczy związany z poczuciem wartości własnej lub aprobatą społeczną (wzorcem autoprezentacji). Wskazywano zwłaszcza, że czynnik ogólny w różnym stopniu nasyca poszczególne sfery osobowości i na ogół jest związany z cechami temperamentu (Zawadzki i Strelau, 2010) oraz może być generowany w wystarczającym stopniu na podstawie neurotyczności i ekstrawersji (Riemann i Kendler, 2010). Ten nurt analiz wynikał z pracy Muska (2007), który generował GFP, także jako czynnik pierwszego rzędu, bezpośrednio z cech PMO, stwierdzając przy tym najwyższe ładunki na wymiarze ogólnym (w kolejności) neurotyczności, ekstrawersji i sumienności, przy zdecydowanie niższych dla ugodowości oraz otwartości. Analizy te podważały zatem faktyczną ogólność GFP. Szczególne znaczenie przypada jednak pracom, w których podnoszona była kwestia zbyt niskiego skorelowania skal PMO, zasadniczo odmiennego od testów zdolności umysłowych, przez co wskazującego na słabość pomiaru czynnika *p*. Sugerowano zatem uwzględnienie innych modeli strukturalnych, np. Spearmanowskiego (czynnik ogólny oraz wymiary specyficzne) w miejsce hierarchicznego (Revelle i Wilt, 2013). Wskazano przy tym, że takie modele analizy danych wskazują na zasadność wyodrębniania czynnika *g* w strukturze skal zdolności umysłowych, przy braku danych potwierdzających niezbędną wprowadzenia czynnika *p* w strukturze skal PMO (Ashton, Lee, Goldberg i De Vries, 2009). Autorzy ci wskazując, że czynnik ogólny nie jest niezbędny do wyjaśnienia skorelowania skal PMO, sugerują też, że to skorelowanie wynika z jednokie-

runkowych związków podwymiarów z czynnikami – w tym sensie jest to problem niedostatecznej konceptualizacji podstawowych cech osobowości w aspekcie dyskryminacyjnym lub niewłaściwego zbalansowania podwymiarów w skali. W tym ostatnim przypadku zatem, gdyby kwestia skorelowania podwymiarów była nie do przezwyciężenia, należałoby uwzględnić przykładowo takie podwymiary neurotyczności, z których jedna część byłaby dodatnio, a druga ujemnie skorelowana z ekstrawersją (oraz podobnie dla ekstrawersji). Przy uzyskaniu odpowiedniego zbalansowania podwymiarów całkowita korelacja między oboma skalami byłaby zerowa i nie zachodziłaby potrzeba generowania żadnej zmiennej ogólnej. Rejestrowane obecnie korelacje między skalami PMO w takim ujęciu są jedynie korelacjami rezydualnymi, a ich niska wartość nie wskazuje na istnienie substancjalnej zmiennej osobowościowej.

Zgadając się z tezą, że wielkość skorelowania wszystkich skal FFM jest kluczowym aspektem wartości diagnostycznej wskaźnika ogólnego, a więc decyduje o jego trafności wewnętrznej oraz trafności w odniesieniu do innych zmiennych psychologicznych, podjęliśmy analizę danych, prezentowaną w tym artykule. O jej odmienności decydują jednak trzy względy. Po pierwsze, dla trafności wewnętrznej każdego czynnika ogólnego istotna jest nie tylko wielkość, lecz także równomierność skorelowania skal składowych. W przeciwnym razie wskaźnik ten będzie miał niskie ładunki czynnikowe dla wszystkich skal lub będzie obejmował tylko niektóre czynniki. Oznacza to, że będzie wykazywał albo słabość pomiarową, albo wadliwą kompozycję treściową – będzie więc miał raczej specyficzny niż ogólny charakter. W tym wypadku analiza powinna jednak zmierzać do wyodrębnienia GFP już jako czynnika pierwszego rzędu w strukturze skal PMO, bez potrzeby analizy hierarchicznej i generowania czynników pośrednich. Po drugie, niskie skorelowanie skal składowych może skutkować zmiennością trafności tego czynnika przy odniesieniu do różnorodnych zmiennych kryterialnych. Innymi słowy, zależnie od wzorca skorelowania skal składowych z tymi zmiennymi kryterialnymi, czynnik ogólny może wykazywać albo wysoką, albo też niską trafność. W tym względzie sytuacja jest zasadniczo odmienna w testach inteligencji, dla których zazwyczaj stwierdza się wysoką trafność czynnika ogólnego *g*, pomimo zróżnicowanej trafności skal składowych (choć nie aż tak różnej, aby uzyskiwać – przy wysokim skorelowaniu skal składowych – zarówno dodatnie, jak i ujemne związki z innymi kryteriami; por. chociażby Brzeziński i Hornowska, 1993; Matczak, Jaworowska, Ciechanowicz i Stańczak, 2006). Po trzecie, z tego względu w prezentowanej analizie uwzględniliśmy szereg różnorodnych zmiennych powiązanych z cechami PMO (objawy PTSD i osobowości schizotypowej, uogólnione poczucie własnej skuteczności, style zachowania w konflikcie oraz behawioralne potencjalne czynniki ryzyka

chorób somatycznych). Są one jednak na tyle odmienne treściowo, aby wykazywać różnokierunkowe powiązania z cechami PMO i umożliwić dokonanie systematycznej analizy trafności GFP.

Hipotezy szczegółowe dotyczące powiązania cech PMO z wybranymi do analizy zmiennymi kryterialnymi zostały przedstawione w innych pracach (zob. Bożko, 2000; Osuch, 2005; Parnowska, 2011). W tym artykule ograniczyliśmy się jedynie do sformułowania oczekiwań odnośnie do:

- słabości pomiaru czynnika ogólnego oraz jego nasycenia zasadniczo poprzez wybrane cechy osobowości;
- zróżnicowanej trafności czynnika GFP w zależności od badanej zmiennej kryterialnej, a więc faktycznie jego ograniczonej wartości diagnostycznej.

METODA

Zaprezentowane dane mają charakter reanalizy wyników, zebranych wcześniej w ramach różnych projektów badawczych. We wszystkich tych badaniach do pomiaru cech PMO zastosowano kwestionariusz NEO-FFI (Costa i McCrae, 1992; Zawadzki i in., 1998). Czynniki GFP wyodrębniono na podstawie sumy standaryzowanych wyników skal PMO na wzór testów inteligencji (po zdekodowaniu wyników w skali Neurotyczności w stronę zrównoważenia emocjonalnego; zob. Rushton i in., 2008) oraz w efekcie zastosowania analizy czynnikowej (metodą Osi Głównych) i zapisania wyniku czynnikowego (zob. Zawadzki i Strelau, 2010). W analizie trafności wyniki skal PMO, jak i czynnika GFP odniesiono do innych zmiennych psychologicznych, takich jak nasilenie objawów PTSD, uogólnione poczucie własnej skuteczności, objawy osobowości schizotypowej, style zachowania w konflikcie oraz behawioralne potencjalne czynniki ryzyka chorób somatycznych¹.

Badane grupy i zastosowane narzędzia badawcze

Pierwsza analiza dotyczyła trafności skal PMO i GFP w stosunku do nasilenia objawów PTSD oraz uogólnionego poczucia własnej skuteczności (por. Kaczmarek i Zawadzki, 2012). W analizie tej nasilenie objawów PTSD badano

¹ W artykule zastosowano pojęcie zmiennej kryterialnej wyłącznie w znaczeniu zmiennej, która posłużyła do weryfikacji trafności wskaźnika GFP. Przedstawione analizy koncentrowały się bowiem na trafności teoretycznej i nie należy ich traktować jako weryfikacji klasycznie rozumianej trafności kryterialnej.

za pomocą kwestionariusza PTSD-C (Strelau, Zawadzki, Oniszczenko i Sobolewski, 2002), zaś uogólnione poczucie własnej skuteczności – polską wersją Skali Uogólnionej Własnej Skuteczności (GSES; Juczyński, 2009; Luszczyńska, Scholz i Schwarzer, 2005). Dane demograficzne tej grupy (oznaczonej schematycznie jako próba 1) oraz pozostałych grup zostały przedstawione w Tabeli 1.

Tabela 1
Charakterystyka demograficzna badanych grup

Próba	<i>N</i>	Płeć	Wiek: zakres	Wiek: <i>M</i> (<i>SD</i>)
1	1132	464 K/668 M	18-66	36,00 (13,57)
2	402	214 K/188 M	18-50	22,42 (4,70)
3	172	101 K/71 M	19-40	21,26 (2,43)
4	200	99 K/101 M	20-25	22,64 (1,46)
Próby 1-4 połączone	1906	878 K/1028 M	18-66	31,43 (13,12)

W drugiej analizie zmienną kryterialną były objawy schizotypowego zaburzenia osobowości. Dane te zostały zebrane przez Parnowską (2011) z zastosowaniem polskiej wersji *Schizotypal Personality Questionnaire* (SPQ), opracowanej przez Rainego (1991). Trzecia analiza dotyczyła stylów zachowania w konflikcie (konsens, unikanie, dominacja, uleganie i kompromis), które zbadano skalą *Rahim Organizational Conflict Inventory* (ROCI II; Rahim, 1983). Polska wersja tego inwentarza została opracowana przez Osuch (2005), która zrealizowała także wykorzystane w tej pracy badania w grupie 3. Ostatnia analiza została przeprowadzona na danych uzyskanych przez Bożko (2000) w próbie 4. Celem tych badań było ustalenie związków między cechami PMO a behawioralnymi potencjalnymi czynnikami ryzyka chorób somatycznych (choroby wieńcowej i nowotworowej). W badaniach wykorzystano inwentarze: Badanie Wzoru Zachowania A (BWZ; Wrześniewski, 1990) do badania wzoru zachowania Typu A, polską wersję Skróconego Inwentarza Reakcji Interpersonalnych (SIRI; Grossarth-Maticzek i Eysenck, 1990), badającego między innymi Typ 1 (podatność na chorobę nowotworową), Typ 2 (podatność na choroby serca) oraz Typ 4 osobowości (osobowość zdrowa – w analizie wykorzystano wyniki skali 4B, ze względu na to, że skala 4A charakteryzuje się znacznie gorszą rzetelnością pomiaru). Ostatnim narzędziem był Inwentarz Wzorów Osobowości (IWO; zob. Zawadzki i Radzikowska, 2006), który bada depresyjność, uległość i wrogość, rozumiane jako osobowościowe potencjalne czynniki ryzyka raka płuca i choroby wieńco-

wej. Po połączeniu tych grup (1-4) uzyskano próbę 1906 osób (878 kobiet i 1028 mężczyzn) w wieku 18-66 lat.

Procedura analizy trafności GFP

Zasadnicza procedura analizy polegała na porównaniu współczynników korelacji ze zmiennymi kryterialnymi, uzyskanych dla skal PMO oraz GFP. Testowanie trafności przyrostowej poprzez analizę regresji nie było możliwe ze względu na współliniowość czynnika GFP w stosunku do wszystkich czy większości skal PMO. Analiza porównawcza korelacji wielorakiej skal PMO faworyzowała natomiast wyniki skalowe w stosunku do korelacji prostej dla GFP ze względu na odpowiednie dopasowanie wag regresji dla zmiennych przewidywanych. Z tego powodu wykonano analizę uproszczoną, polegającą na porównaniu trafności pojedynczych skal PMO najwyższej skorelowanych z daną zmienną kryterialną (w Tabeli 3 zaznaczono je pogrubioną czcionką) z odpowiednią korelacją uzyskaną dla czynnika GFP. W tym wypadku testowano różnicę obu korelacji, uwzględniającą skorelowanie danej skali PMO i GFP dwustronnym testem t (zob. Cohen i Cohen, 1983). Formalna analiza trafności przyrostowej dla jednej skali PMO i GFP także nie była możliwa lub niemiarodajna ze względu na uzyskiwanie: (1) zbyt wysokiej współliniowości obu zmiennych predykcyjnych, zwłaszcza dla neurotyczności, (2) przy wielocechowym uwarunkowaniu zmiennych kryterialnych – istotnych wartości przyrostowych wyjaśnionej wariacji zarówno dla GFP, jak i skali PMO o najwyższej trafności, oraz (3) istotne korelacje semicząstkowe dla GFP, przy jego niskich korelacjach prostych ze zmienną kryterialną albo wręcz zmiany znaku tego współczynnika w stosunku do korelacji prostej, zwłaszcza w przypadku ugodowości. Może to sugerować występowanie efektu przypadkowej optymalizacji rozwiązania w próbie. Wynik taki byłby zresztą trudny do teoretycznej interpretacji. Z tego względu analizę porównawczą trafności ograniczono do prostego porównania korelacji skali PMO oraz czynnika GFP ze zmiennymi kryterialnymi.

WYNIKI

Trafność strukturalna wskaźnika GFP

Analiza danych została rozpoczęta od wyodrębnienia Ogólnego Czynnika Osobowości dla każdej próby oraz wszystkich prób połączonych. W Tabeli 2 przedstawiono ładunki czynnikowe skal NEO-FFI dla czynnika ogólnego, jego

wartości własne, medianę korelacji skal PMO oraz współczynnik α Cronbacha, ilustrujący rzetelność pomiaru tej zmiennej (obliczony dla wyników pięciu skal).

Tabela 2

Ładunki czynnikowe oraz wagi β skal PMO (z NEO-FFI) dla GFP

Ładunki czynnikowe skal PMO	Próba 1	Próba 2	Próba 3	Próba 4	Próby 1-4 połączone
N	-0,68	-0,61	-0,69	-0,68	-0,65
E	0,59	0,69	0,67	0,68	0,63
O	0,08	0,28	0,30	0,11	0,14
A	0,37	0,39	0,33	0,30	0,36
C	0,54	0,56	0,40	0,63	0,52
Wartość własna	1,23 (24,6%)	1,39 (27,8%)	1,29 (25,7%)	1,42 (28,5%)	1,23 (24,6%)
Korelacje między skalami (zakres)	<i>Mdn</i> = 0,19 (-0,03 do 0,40)	<i>Mdn</i> = 0,25 (0,06 do 0,42)	<i>Mdn</i> = 0,20 (0,11 do 0,50)	<i>Mdn</i> = 0,20 (-0,02 do 0,46)	<i>Mdn</i> = 0,20 (-0,03 do 0,41)
α Cronbacha	0,54	0,62	0,59	0,59	0,56
Wagi β skal PMO	Próba 1	Próba 2	Próba 3	Próba 4	Próby 1-4 połączone
N	-0,52	-0,36	-0,50	-0,43	-0,47
E	0,38	0,50	0,47	0,43	0,44
O	0,04	0,10	0,13	0,04	0,06
A	0,18	0,17	0,15	0,12	0,18
C	0,32	0,30	0,19	0,35	0,30

Uwaga. N – Neurotyczność, E – Ekstrawersja, O – Otwartość na doświadczenie, A – Ugodowość, C – Sumienność. Przy obliczaniu współczynnika α Cronbacha oraz sumy skal PMO wyniki skali N zostały zdekodowane w stronę zrównoważenia emocjonalnego. Różnice współczynnika rzetelności pomiaru między badanymi próbami były nieistotne statystycznie ($\chi^2 = 4,90$ dla $df = 3$, podobnie jak różnice między grupami jednorodnymi ze względu na płeć wyodrębnionymi w próbach połączonych 1-4: $\chi^2 = 0,11$ dla $df = 1$ oraz trzema grupami wieku: $\chi^2 = 0,11$ dla $df = 2$; zob. Hakstian i Whalen, 1976). Współczynniki rzetelności skal NEO-FFI dla prób połączonych (1-4): 0,84 (N), 0,78 (E), 0,69 (O), 0,68 (A) i 0,82 (C). Czynniki ogólny (GFP) wyodrębniono metodą Osi Głównych dla wyników standaryzowanych skal. Dla sumy skal wagi β były równe 0,34 (Próba 1), 0,32 (Próba 2) oraz 0,33 (Próba 3 i 4 oraz próby połączone) dla wszystkich skal NEO-FFI. Współczynniki Tuckera podobieństwa czynnikowego w poszczególnych grupach w stosunku do rozwiązania uzyskanego w próbie połączonej: od 0,982 do 0,997, przy porównaniach poszczególnych prób najniższy współczynnik uzyskano dla grup 1 i 3 (0,970).

Uzyskane wyniki są zgodne z oczekiwaniami. Po pierwsze, dane wskazują, że ogólny czynnik osobowości jest nasycony tylko niektórymi cechami osobowości – zasadniczo: neurotycznością i ekstrawersją, w mniejszym stopniu sumiennością, zaś nisko – ugodowością, a szczególnie otwartością. Struktura czynnikowa była przy tym bardzo spójna dla różnych grup, na co wskazuje wysoka wartość współczynników podobieństwa czynnikowego. Dane te w pełni potwierdzają ustalenia innych naszych analiz, w których wykazaliśmy, że GFP faktycznie redukuje się do cech temperamentu, a ponadto efekt ten nie wydaje się wynikać ze zróżnicowania związanego z płcią i wiekiem osób badanych, bowiem uzyskane dane są bardzo zbliżone do wyników otrzymanych dla danych skorygowanych ze względu na te zmienne demograficzne (Zawadzki i Strelau, 2010). Po drugie, zarówno współczynniki rzetelności pomiaru (o wartości koło 0,60), jak i wielkość wyjaśnionej wariancji skal (rzędu około 25%) wskazują na słabość pomiaru zmiennej ogólnej. Zasadniczo wynika to z niskiego skorelowania skal NEO-FFI (mediana korelacji w badanych grupach rzędu 0,20). W efekcie także próba wyjaśnienia zmienności cech PMO poprzez jedną wspólną zmienną latentną nie znalazła dostatecznego wsparcia empirycznego – model z jednym czynnikiem wyższego rzędu w analizie konfirmacyjnej wyników skal PMO nie wykazał dobrego dopasowania do danych (w próbach połączonych 1-4): $\chi^2 = 86,40$, $df = 5$, $p = 0,00$, $RMSEA = 0,091$, $GFI = 0,993$. W analizach walidacyjnych – pomimo ewidentnych wad pomiaru czynnika ogólnego – uwzględniono przy tym dwa wskaźniki GFP: czynnikowy (pochodny od analizy eksploracyjnej) oraz sumaryczny (suma standaryzowanych wyników skal NEO-FFI). Ten ostatni wskaźnik, sugerowany przez Rushtona i innych (2008) – na wzór generowania wyniku ogólnego w testach inteligencji, był jednak traktowany wyłącznie jako dodatkowy. Dla obu wskaźników GFP wyznaczono też wagi β , ilustrujące wkład danej skali PMO do GFP, w efekcie zastosowania analizy regresji ze skalami NEO-FFI jako zmiennymi niezależnymi dla GFP.

Trafność wskaźnika GFP ***– analiza związków ze zmiennymi kryterialnymi***

W Tabeli 3 przedstawione zostały współczynniki korelacji skal PMO oraz wskaźników (czynnikowego i sumarycznego) GFP ze wszystkimi zmiennymi kryterialnymi. Wskazano także skale PMO, które były predyktorami zmiennych kryterialnych w analizie regresji (nie uwzględniającej GFP) oraz wyniki porównania korelacji kryterialnych dla wskaźnika czynnikowego GFP i skali PMO o najwyższej trafności.

Tabela 3

Trafność skal PMO oraz czynnika GFP – korelacje z objawami PTSD, osobowości schizotypowej, stylami zachowania w konflikcie oraz behawioralnymi potencjalnymi czynnikami ryzyka chorób somatycznych (choroby wieńcowej i nowotworowej)

Skale PMO	Próba 1: PTSD	Próba 1: Poczucie skuteczności	Próba 2: Osobowość schizotypowa	Próba 3: Konsens	Próba 3: Unikanie
N	0,44*	-0,49*	0,48*	-0,34*	0,31*
E	-0,18*	0,39*	-0,35*	0,34*	-0,22*
O	0,01	0,07*	0,01	0,33*	-0,24*
A	-0,05*	0,05	-0,30*	0,39*	0,05
C	-0,12*	0,36*	-0,30*	0,27*	0,03
GFP	-0,34*	0,53*	-0,49*	0,48*	-0,27*
Suma skal	-0,26*	0,46*	-0,45*	0,54*	-0,22*
Predyktory	N	nEoAC	NeOa	nEOA	No
GFP vs PMO	PMO (0,01)	GFP (0,01)	n.i.	n.i.	n.i.
Skale PMO	Próba 3: Dominacja	Próba 3: Uleganie	Próba 3: Kompromis	Próba 4: Typ 1	Próba 4: Typ 2
N	-0,08	0,18*	-0,03	0,37*	0,38*
E	0,16*	-0,12	0,17*	-0,33*	-0,22*
O	0,05	-0,16*	0,10	-0,23*	-0,06
A	-0,44*	-0,11	0,29*	-0,13	-0,16*
C	0,03	-0,12	0,06	-0,18*	-0,09
GFP	0,06	-0,17*	0,16*	-0,36*	-0,31*
Suma skal	-0,04	-0,15*	0,21*	-0,32*	-0,30*
Predyktory	eA	NoA	A	NeoA	N
GFP vs PMO	PMO (0,01)	n.i.	n.i.	n.i.	n.i.
Skale PMO	Próba 4: Typ 4	Próba 4: Typ A	Próba 4: Depresyjność	Próba 4: Wrogość	Próba 4: Uległość
N	-0,55*	0,07	0,75*	0,32*	0,13*
E	0,41*	0,33*	-0,56*	0,03	-0,12*
O	0,13	0,15*	-0,01	0,06	-0,23*
A	0,22*	-0,40*	-0,08	-0,45*	0,32*
C	0,31*	0,08	-0,38*	-0,14*	0,02
GFP	0,55*	0,10	-0,70*	-0,22*	-0,07
Suma skal	0,53*	0,03	-0,58*	-0,27*	-0,05
Predyktory	nE	NEa	Ne	NEa	NoA
GFP vs PMO	n.i.	PMO (0,01)	n.i.	PMO (0,01)	PMO (0,01)

Uwaga. Predyktory – skale PMO wykazujące istotne korelacje semicząstkowe w analizie regresji ze zmiennymi kryterialnymi (bez uwzględnienia GFP). * – współczynniki korelacji istotne na $p < 0,05$. Najwyższe korelacje skal PMO zostały oznaczone pogrubioną czcionką, podobnie jak korelacje czynnika GFP. Dla porównania korelacji ze zmiennymi kryterialnymi między GFP oraz jedną ze skal PMO zastosowano test t (dwustronny; zob. Cohen i Cohen, 1983).

Jedynie trzy (na 15) analizowane współczynniki korelacji wskaźnika GFP ze zmiennymi kryterialnymi nie były istotne statystycznie, a ich mediana wyniosła 0,34 (przy względnie dużym podobieństwie danych uzyskanych dla wskaźnika czynnikowego i sumarycznego). Wynik ten pozornie jednak wskazuje na zadowalającą trafność GFP. Pełny dowód wymaga analizy uwzględniającej także trafność skal źródłowych oraz bliższej analizy przypadków wskazujących na wysoką, jak i niską trafność GFP. W analizie zarejestrowano bowiem również znaczny rozrzut współczynników korelacji dla GFP: od 0,06 do 0,70, co już samo w sobie jest zgodne ze sformułowanymi oczekiwaniami.

Jak sugerowano w części teoretycznej, wskaźnik GFP będzie miał wysoką trafność, gdy (1) skale źródłowe PMO będą wykazywały wysokie korelacje ze zmiennymi kryterialnymi oraz (2) wielkość i kierunek tych korelacji będzie zbliżony z ładunkami czynnikowymi czy wartościami β , przy wyodrębnianiu GFP. Powinien zatem odpowiadać profilom cech (lub w niewielkim stopniu od nich odbiegać), kodowanym schematycznie jako „nEOAC” czy jego odwrotność, a zatem „Neoac”. Przykładem są tu wysokie korelacje, zarówno dla cech PMO, jak i GFP, uzyskane dla konsensualnego stylu zachowania w konflikcie (profil: „nEOA” w analizie regresji), uogólnionego poczucia własnej skuteczności („nEoAC”), objawów schizotypii („NeOa”), Typu 1 osobowości („nEoA”), Typu 4 osobowości („nE”) czy depresyjności („Ne”); skale o wzorcu skorelowania odwrotnym do profili, stanowiących opisywane bieguny GFP, wyróżniono pismem półgrubym). Odwrotna korelacja, uzyskana dla ugodowości czy otwartości, nie stanowi przy tym istotnego problemu, pod warunkiem, że korelacje pozostałych cech są zgodne ze wzorcem i odpowiednio wysokie. Nie prowadzi ona zatem do obniżenia trafności GFP, zasadniczo z tego powodu, że wkład tych skal do wyniku czynnikowego (zwłaszcza otwartości) jest marginalny (zob. wartości β w Tabeli 2). Należy zauważyć jednak, że tylko w jednym przypadku (uogólnione poczucie własnej skuteczności) korelacja GFP jest istotnie wyższa od najwyższej korelacji uzyskanej dla skal PMO, zaś w innych trafność pojedynczej skali PMO jest praktycznie identyczna z trafnością wskaźnika GFP.

Pozostałe przypadki dotyczą jednak już zdecydowanie niższej trafności GFP. Relatywnie niskie współczynniki korelacji uzyskano aż w trzech odmiennych sytuacjach. Pierwsza z nich odnosi się do przypadku stosunkowo niskiej trafności samych skal PMO, co stwierdzono dla stylu zachowania w konflikcie, polegającego na unikaniu lub uleganiu. Przyczyna niskiej trafności GFP jest tu jednak oczywista – przy braku trafności skal źródłowych ich wynik zagregowany nie może być trafny. Sytuacja druga odnosi się do małej trafności GFP, wynikającej ze zróżnicowanej trafności poszczególnych skal PMO, a więc gdy tylko

niektóre z nich wykazują odpowiednio wysoką trafność (w sensie teoretycznym, gdy dana zmienna kryterialna wykazuje powiązanie tylko z jedną cechą PMO). Dotyczy to zarówno przypadku, gdy dana skala NEO-FFI ma duży wkład do GFP (wysoką wagę β), na przykład, jak stwierdzono, dla nasilenia objawów PTSD czy Typu 2 osobowości („N”), a zwłaszcza gdy wkład ten jest relatywnie mały (niska waga β) – styl kompromisowy zachowania w konflikcie (tylko „A”). Przyczyna niższej trafności GFP wydaje się wynikać z „rozcieńczenia” trafności danej skali PMO – występującego, gdy jest ona agregowana w wynik ogólny ze skalami, które tej trafności nie wykazują. Najniższe współczynniki korelacji uzyskano dla GFP w ostatniej sytuacji, mianowicie gdy profil korelacji znacznie odbiegał od profilu „nEOAC” czy „Neoac”, co stwierdzono dla dominacyjnego stylu zachowania w konflikcie („Ea”), Typu A osobowości („NEa”), wrogości („NEa”) czy uległości („NoA”). Dla wszystkich tych zmiennych stwierdzono przy tym stosunkowo wysokie korelacje skal PMO (istotnie wyższe niż korelacja GFP). Przypadek ten jest przy tym dowodowo najważniejszy, bowiem wskazuje na niską trafność wskaźnika GFP, przy wysokiej trafności poszczególnych skal źródłowych. W tym wypadku następuje jednak nie tyle „rozcieńczenie” trafności, ile zniwelowanie trafności czynnika ogólnego wskutek znoszenia się trafności poszczególnych skal przy ich agregowaniu w wynik ogólny. Ma ono miejsce, gdy wzór korelacji z daną zmienną kryterialną jest odmienny od wzorca agregowania w GFP, np. duże znaczenie wykazuje zarówno wysoka neurotyczność, jak i wysoka ekstrawersja. Przy agregowaniu wyników tych skal ich wkład do trafności GFP wręcz się znosi, pozbawiając go znaczenia empirycznego, pomimo zachowanej trafności skal źródłowych. Stosunkowo niskie korelacje skal PMO pozwalają przy tym sądzić, że w badaniach osobowości sytuacja taka wcale nie byłaby rzadka, a zatem dla wielu zmiennych kryterialnych możliwe byłoby uzyskanie danych wskazujących na wysoką trafność skal PMO, przy niskiej lub wręcz „zerowej” trafności GFP.

DYSKUSJA

Uzyskane wyniki potwierdziły oba sformułowane oczekiwania badawcze. Przede wszystkim stwierdzono znaczny rozrzut współczynników korelacji czynnika GFP ze zmiennymi kryterialnymi. Już sam ten fakt dowodzi stosunkowo małej przydatności diagnostycznej GFP. Zrealizowana analiza pozwala jednak także na wyjaśnienie przyczyn tej zmienności, a mianowicie wskazuje, jakie warunki zostały spełnione, gdy czynnik ogólny charakteryzowała wysoka traf-

ność, oraz w jaki sposób niespełnienie tych warunków rzutowało na obniżenie trafności. Zasadniczo dotyczą one stosunkowo oczywistego wymogu trafności skal źródłowych oraz zachowania zbieżności kierunku i wielkości korelacji walidacyjnych z wartościami β otrzymanymi przy wyodrębnianiu GFP. Szczególne znaczenie mają przy tym dane wskazujące na wysoką trafność skal źródłowych przy niskiej trafności czynnika ogólnego. W tych przypadkach bowiem agregacja wyników skal PMO, odwołująca się do ich wewnętrznej struktury i odmienna od wzoru predykcji zmiennych kryterialnych, prowadzi do znoszenia się trafności cech źródłowych i obniżania się trafności czynnika ogólnego. Przyczyna nietrafności, jak i zresztą trafności czynnika ogólnego tkwi zatem w samym akcie agregacji, jeśli jest on niejako „zewnętrzny” w stosunku do zmiennych kryterialnych. Efekt ten nie występuje w analizie regresji, uwzględniającej możliwość trafności zarówno wysokiego, jak i niskiego nasilenia danej cechy dla danej zmiennej przewidywanej. Czy zatem wyniki te przemawiają generalnie za odrzuceniem agregowania wyników? W jakimś sensie bowiem dane zagregowane zawsze wiążą się z utratą informacji o specyfice zachowania. W psychologii różnic indywidualnych jednak agregacja jest powszechna – wyniki pozycji są łączone w skale, zaś podwymiary – w cechę ogólną, np. wyniki testów zdolności umysłowych – w inteligencję ogólną, zaś podwymiary, takie jak depresja, lęk, nadwrażliwość itp., w cechę ogólną osobowości, jaką jest neurotyczność. Jaka agregacja zatem ma sens, a dokładniej, na jakim poziomie agregacji należałoby się zatrzymać? Kryterium jest stosunkowo proste i wymaga odpowiednio wysokiego skorelowania agregowanych zmiennych oraz zatrzymania analizy na poziomie agregacji, gdy korelacje zmiennych wyższego rzędu są stosunkowo niskie lub najlepiej – zerowe (Nowakowska, 1975). W przypadku wysokich korelacji skal źródłowych mamy bowiem przesłanki do wniosku, że badana jest nowa zmienna, niesprowadzalna wyłącznie do zmiennych źródłowych, a trafność wskaźnika ogólnego powinna być stabilnie wysoka przy odniesieniu do różnych zmiennych kryterialnych. Przy wysokim skorelowaniu skal bardzo rzadkie będą także przypadki odwrotnej trafności skal źródłowych w stosunku do klucza agregacji, skutkujące obniżaniem się trafności czynnika ogólnego. Idea tworzenia czynnika GFP w oczywisty sposób narusza tę prostą regułę – korelacje skal PMO są zbyt niskie, aby sensownie wnioskować o czynniku ogólnym, a wymuszona agregacja skutkuje znacznym rozrzutem trafności tego czynnika. Wyniki wskazujące zaś na dużą trafność GFP w zasadzie niczego nie rozstrzygają – arbitralna suma ortogonalnych skal także wykazywałaby dużą trafność, pod warunkiem wysokiej trafności skal źródłowych. Rzecz w tym bowiem, że natychmiast traciłaby trafność przy uwzględnieniu innej zmiennej kryterialnej. Dla każdej z takich zmiennych

można byłoby tworzyć każdorazowo odmienne agregaty skal źródłowych, charakteryzujących się wysoką trafnością wyłącznie dla przewidywania tej zmiennej. W bardziej ogólnej perspektywie pozostawalibyśmy wówczas wyłącznie na poziomie cech źródłowych, łączonych w odpowiednie kombinacje, zależnie od potrzeb badawczych. A skoro faktycznie pozostawalibyśmy na poziomie cech źródłowych, to sens tworzenia ich arbitralnych agregatów byłby mocno wątpliwy.

Podstawowa dla tych rozważań jest zatem kwestia, czy poziom skorelowania skal PMO jest wystarczający, aby ich agregacja miała sens, czy – patrząc z innej perspektywy – GFP odzwierciedla ważną zmienną osobowościową (przyjmując, że cechy PMO pozwalają na wyczerpującą charakterystykę osobowości). Przede wszystkim należy stwierdzić, że skorelowanie skal badających cechy PMO jest faktem. Dobrze zostało to podsumowane w leksykalnych badaniach nad osobowością przez Szarotę (1995, s. 248), który stwierdza wręcz, że „[...] istotne korelacje między wymiarami Big Five, relacjonowane były już tak często, że w ich ortogonalność przestali już wierzyć nawet sami teoretycy”. Z drugiej jednak strony uważa się, że korelacje te są zbyt niskie, aby mogły wskazywać na pomiar istotnej zmiennej psychologicznej, zwłaszcza przy porównaniu z testami inteligencji (Ashton i in., 2009; Revelle i Wilt, 2013). Wyniki takie uzyskano także w niniejszym artykule i wskazują one na zasadniczą słabość pomiaru GFP. Uważa się przy tym, że stanowią one korelacje resztowe, wynikające z jednokierunkowego powiązania różnych podwymiarów z cechami ogólnymi (głównie neurotyczności i ekstrawersji; por. Riemann i Kandler, 2010), i nie wskazują na istnienie ważnych zmiennych wyższego rzędu (Ashton i in., 2009). Wyniki analiz Muska (2007), Riemanna i Kandlera (2010), jak i naszych badań, przedstawionych w innej pracy (Zawadzki i Strelau, 2010), wskazują przy tym na nierównomierne nasycenie czynnikiem ogólnym poszczególnych cech PMO. Takie same dane uzyskano także w tym artykule; jak się okazuje, czynnik ogólny wcale nie ma ogólnego charakteru i zasadniczo wyznaczają go cechy natury temperamentalnej (neurotyczność i ekstrawersja). W sensie teoretycznym jest to ważny argument na rzecz tezy o osiowej roli cech temperamentu w strukturze osobowości. Na tym jednak – jak sądzimy – waga teoretyczna analiz GFP praktycznie się kończy. Tworzenie agregatów zmiennych ma bowiem zawsze ograniczony sens empiryczny, jeśli nie stoją za tym ani dostateczne argumenty teoretyczne, ani też metodologiczne (zasadniczo strukturalne), uzasadniające tworzenie takich agregatów.

LITERATURA CYTOWANA

- Ashton, M. C., Lee, K., Goldberg, L. R. i 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*. Niepublikowana praca magisterska, Wydział Psychologii UW, Warszawa.
- Brzeziński, J. i Hornowska, E. (red.) (1993). *Skala inteligencji Wechslera. WAIS-R. Polska adaptacja, standaryzacja, normalizacja i wykorzystanie w diagnostyce psychologicznej*. Warszawa: Wydawnictwo Naukowe PWN.
- Cohen, J. i Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Costa, P. T., Jr. i 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. i 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-Maticzek, R. i Eysenck, H. J. (1990). Personality, stress and disease: Description and validation of a new inventory. *Psychological Reports*, 66, 355-373.
- Hakstian, A. R. i Whalen, T. E. (1976). A *k*-sample significance test for independent alpha coefficients. *Psychometrika*, 41, 219-231.
- Holden, R. R. i 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*. Warszawa: 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. i 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. i Schwarzer, R. (2005). The General Self-Efficacy Scale: Multicultural validation studies. *The Journal of Psychology*, 139, 439-457.
- Matczak, A., Jaworowska, A., Ciechanowicz, A. i Stańczak, J. (2006). *Bateria testów APIS-Z*. Warszawa: 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*. Warszawa: PWN.
- Osuch, K. (2005). *Style zachowania w konflikcie a cechy osobowości*. Niepublikowana praca magisterska, Wydział Psychologii UW, Warszawa.
- Parnowska, D. (2011). *Genetyczne i środowiskowe uwarunkowania struktury schizotypii u osób zdrowych*. Niepublikowana praca doktorska, Instytut Psychiatrii i Neurologii, Warszawa.
- 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. i Wilt, J. (2013). The general factor of personality: A general critique. *Journal of Research in Personality*, 47, 493-504.
- Riemann, R. i 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. i 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. i Irving, 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. i Irving, 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. i 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. i 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. i Radzikowska, E. (2006). Próba kwestionariuszowej diagnozy potencjalnych osobowościowych czynników ryzyka raka płuca i choroby wieńcowej. *Psychologia – Etologia – Genetyka*, 13, 135-154.
- Zawadzki, B. i 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. i Śliwińska, M. (1998). *Inwentarz Osobowości NEO-FFI Costy i McCrae – adaptacja polska. Podręcznik*. Warszawa: Pracownia Testów Psychologicznych PTP.