

Właściwości skali a wynik pomiaru kwestionariuszowego

Katarzyna Trawka

Instytut Psychologii, Uniwersytet Opolski, Opole

Tomasz Wirga

Instytut Psychologii, Uniwersytet Opolski, Opole

Romuald Derbis*

Instytut Filozofii, Socjologii i Psychologii, Akademia im. J. Długosza, Częstochowa

PROPERTIES OF THE SCALES AND THEIR INFLUENT ON THE RESULTS IN QUESTIONNAIRE SURVEYS

The purpose of this article is to present what impact to the results of a questionnaire research have the formal aspects of the scales: negative numeric values, zero-start point, kind of labels and length of numeric scale. First research shown, that change of numeric value from positive to negative, without changing the length, drives to crystallization of answers. The middle neutral point „no opinion” is chosen more rarely, than options, which are placed to the right from the midpoint. Answer options placed to the left from midpoint, were chosen most often. In case of „positive” scales, with zero-start point and one-start point, changes of labels do not differentiate the results. In case of „negative” scale, labels, which describe only the end points of scale, will cause left side shift of the results as against to results obtained by „every point” labels. Second research was conducted to verify if the results given in first study are characteristic only to our questionnaire or if they are more versatile. The deeper exploration of the given result was made on NEO-FFI questionnaire. In general, effects were confirmed. In this article, results are also analyzed in methodological and ethical context.

WPROWADZENIE

W badaniach kwestionariuszowych opartych na samoopisie wyniki uzyskane przez badacza zależne są od wielu czynników podmiotowych, przedmiotowych oraz ich interakcji. Czynniki podmiotowe dotyczą osoby badanej i badacza. Związane są z ich etyką, motywacją do udziału i prowadzenia badań, potrzebami, postawami, lękiem przed oceną, cechami osobowości i temperamentu oraz doświadczeniami kulturowymi. Czynniki przedmiotowe przede wszystkim związane są z właściwościami i struk-

turą pytań tworzących skale, a także z ukrytymi wymaganiami sytuacji badawczej oraz instrukcjami maskującymi. Na uzyskany wynik wpływ może mieć forma twierdzeń lub pytań, liczba pozycji, liczba kategorii możliwych odpowiedzi, sposób opisu etykiet na skali, jasność języka, stopień drażliwości przedmiotu pomiaru.

W polskiej literaturze od dawna problemy te podejmowane są *explicite* w obrębie tematyki metodologicznej i psychometrycznej. Egzemplifikacją mogą być opracowania autorskie lub redakcyjne Nowakowskiej (1975, 1977), Brzezińskiego (1978, 1978a, 1984, 1993, 1996, 2000), Brzezińskiego i Hornowskiej (1993), Brzezińskiego i Siuty (1991), Brzezińskiego i Toeplitz-Winiewskiej (2000), Paluchowskiego (1983, 1991), Drwała (1989, 1995) czy Aranowskiej (1996). Brzeziński (1993a) wprowadza

* Korespondencję dotyczącą artykułu można kierować na adres: Romuald Derbis, Instytut Filozofii, Socjologii i Psychologii Akademii im. J. Długosza, ul. Waszyngtona 4/8, 42-200 Częstochowa.
e-mail: rderbis@neostrada.pl

przydatne pojęcie przestrzeni diagnostycznej, która ograniczona jest wymiarami treściowo związanymi z badaczem i badanym. W szeregu opracowań z nauk społecznych opartych na danych samoopisowych, których tu nie sposób wymienić personalnie, pojawia się *implicite* problem zmienności wyników w zależności od wpływu wyżej wymienionych czynników.

Począwszy od eksperymentów laboratoryjnych, przez badanie opinii publicznej, aż do działań telemarketingowych, badacze polegają na odpowiedziach udzielanych przez uczestników badań. Wyciągają na tej podstawie wnioski, co do ich myśli, uczuć i zachowań. Informacje uzyskane w ten sposób służą także wskazywaniu aktualnych trendów społecznych, takich jak skuteczność programów dla bezrobotnych czy rozwój przestępczości w kraju. Schwarz (1999) zwraca jednak uwagę, że nawet najmniejsza zmiana w słownictwie użytym w pytaniu, rodzaj pytania czy nawet kontekst oraz kolejność pytań znacząco wpływa na zmianę uzyskiwanych wyników. Metoda samoopisowa może więc być bardzo labilnym źródłem informacji.

Jednymi z najczęściej zadawanych przez psychologów pytań w kwestionariuszach są te z dołączoną do nich skalą numeryczną. Przypuśćmy, że jest to pytanie „Jak oceniasz miniony rok?”, któremu towarzyszy skala od „całkowicie nieudany” do „wyjątkowo udany”. Osoba badana, aby odpowiedzieć na to pytanie stara się określić, co badacz rozumie przez „całkowicie nieudany” – czy chodzi mu o nieobecność wyjątkowo sprzyjających wydarzeń i sukcesów, czy też o wystąpienie widocznych potknięć i porażek, a może o jedno i drugie? Aby odpowiedzieć sobie na to pytanie, respondent może posłużyć się czysto formalnym elementem pytania – skalą numeryczną, a konkretnie jej wartościami liczbowymi. Schwarz, Knäuper, Hippler i Noelle-Neuman (1991) zaprezentowali swoim badanym pytania dotyczące sukcesów w życiu zaopatrzone w 11 punktową skalę, która w zależności od warunków badawczych rozpoczynała się od 0 (brak sukcesów) do 10 (wyjątkowe osiągnięcia) lub od -5 (brak sukcesów) do +5 (wyjątkowe osiągnięcia). Otrzymane wyniki pokazują, że 34% badanych umieściło swoją odpowiedź w przedziale -5 do 0 gdy zastosowano skalę od -5 do +5, podczas gdy jedynie 13% osób swoją odpowiedź zaznaczyło w przedziale 0 do 5 gdy zastosowano skalę od 0 do 10. Autorzy badań sugerują, że za tak dużą rozbieżność w otrzymanych wynikach mogą odpowiadać różnice w interpretacji terminu „brak sukcesów”. Kiedy termin ten był połączony z wartością skali równą „0” badani mogli to zinterpretować jako wskaźnik braku oszała-

miających sukcesów. Jednakże, kiedy ten sam termin połączono z wartością skali równą „-5” i skala oferowała także wartość „0” jako punkt środkowy, mogli interpretować ten termin jako występowanie jedynie drobnych porażek (por. Schwarz i Hippler, 1995). Ogólnie mówiąc, skala, która zaczyna się od wartości ujemnych a kończy na dodatnich komunikuje, że badacz ma na myśli dwubiegunowy wymiar, którego krańce odnoszą się do występowania przeciwnych cech. Z kolei, skala, która wykorzystuje tylko liczby dodatnie komunikuje, że badacz ma na myśli jeden wymiar dotyczący natężenia danej cechy (Schwarz, 1991).

Dwubiegunowość skal ujemno-dodatnich wydaje się być zgodna ze zdrowym rozsądkiem i logiczna. Jednakże z rozumieniem przez badanego idei skali tylko dodatniej mogą wiązać się trudności. Gdy skala rozpoczyna się od „1” sprawa jest względnie jasna – ten koniec skali dotyczy bardzo małego natężenia analizowanej cechy. Trudność pojawia się gdy skala rozpoczyna się od „0”. Czy badany ma „0” rozumieć jako brak danej cechy, czy może także jako jej bardzo niski poziom? Problem leży w indywidualnej mentalnej reprezentacji zera, wynikającej z jego specyfiki.

Najstarsze znane dokumenty zawierające znaki liczb sięgają czwartego tysiąclecia p.n.e. Zastanawiające jest bardzo późne powstanie samego znaku zero. Zero jako liczba zostało wprowadzone do powszechnego użytku przez Hindusów dopiero w średniowieczu, podczas gdy pozostałe liczby były już wykorzystywane przez Babilończyków około V w. p.n.e. (Kordos, 1994). Jak i czym wytłumaczyć, że jeden ze znaków, za pomocą których zapisujemy liczby, powstał kilka tysięcy lat później od innych? Fakt ten możemy wyjaśnić trudnością uznania zera za liczbę. Zero jako odpowiednik wyrazu „nic” wydaje się czymś innym od innych liczb (Krysicki, 1973).

Gilbert A. Churchill Jr (2002), w swojej książce o podstawach metodologicznych badań marketingowych, a dokładniej w rozdziale poświęconym konstruowaniu kwestionariuszy do zbierania danych przytacza kilka bardzo ciekawych wyników badań nad wpływem własności skali na uzyskiwane odpowiedzi. Podaje on wyniki pewnej analizy intencji dokonywania zakupów, w której używano zarówno czteropunktowej, jak i pięciopunktowej skali. Skala pięciopunktowa różniła się od czteropunktowej jedynie środkową wartością neutralną. Podsumowując wyniki tych badań, można wysunąć wniosek, że jeżeli weźmiemy pod uwagę jedynie skrajne wartości skali („zdecydowanie kupię” i „zdecydowanie nie kupię”) to dodatkowy punkt na skali nie wpływa

na rozkład odpowiedzi. Jeżeli natomiast weźmiemy pod uwagę całą skalę, to okazuje się, że skala bez neutralnego punktu dostarcza więcej pozytywnych odpowiedzi niż skala rozszerzona.

Inni badacze (np. Schwarz i in, 1985; Greenleaf, 1992) ilustrują kolejny problem powstający w związku z wieloma wariantami odpowiedzi na temat częstości różnych zjawisk. Skala kategorii używanych w pytaniu zdaje się umożliwiać osobom badanym podanie właściwej odpowiedzi. Jednakże jak podają wyniki badań kategorii odpowiedzi z góry wpływają na obraz częstości rozpatrywanych odpowiedzi. Wnioskować można, że respondenci niejako dostosowują się do podanych oczekiwań, tzn. niechętnie ujawniają zachowania, które wg podanej skali są jakies niezwykle, czyli te, które tworzą kategorie skrajne.

Czapiński (1998) przeprowadził badania mierzące zmienność jakości życia w czasie. Analizował szczegółowo, jaki udział w kształtowaniu dobrostanu psychicznego mają czynniki socjodemograficzne w stosunku do czynników psychologicznych. Posługiwał się tu dwoma terminami: „uziemienie polskiej duszy” i „lewitacja”. Przez uziemienie Czapiński rozumie znaczne uwarunkowanie dobrostanu (np.: depresji) twardymi wyznacznikami społecznymi (np.: płeć czy wykształcenie). Lewitacją określa znaczny udział czynników psychologicznych. Czapiński zaobserwował, że w pierwszych trudnych latach transformacji (1989 – 1990) społeczeństwo polskie bardziej lewitowało, niż w latach późniejszych (1995-1996), gdzie było zdecydowanie bardziej uziemione. Jednakże, gdy zmienił 7-punktową miarę zadowolenia z życia na 4-punktową, otrzymane wyniki zmieniły się. Społeczeństwo polskie wydawało się lewitować również w latach 1988-1994, ale według wyników na skali 4-punktowej oznaki uziemienia pojawiły się już w następnym roku. Pozostało więc bez jasnej odpowiedzi pytanie o to, czy uziemienie jest uwarunkowane historycznie, czyli transformacją, czy też kulturowo, czyli tradycją? Mówiąc inaczej nie wiemy, co było prawdziwą przyczyną uziemienia duszy polskiej – czy uwarunkowania kulturowe, historyczne czy też właściwości prezentowanej badanym skali.

Poszerzenie naszej wiedzy na temat metod samoopisowych w celu doprowadzenie do podniesienia jakości zbieranych informacji wymaga od badaczy szukania odpowiedzi na szereg ważkich pytań:

- Jak respondent może zrozumieć zadane mu pytanie?
- Jakie czynniki, oprócz treści pytań, wpływają na rozkład odpowiedzi?

- Jakie znaczenie ma pamięć autobiograficzna w pytaniach dotyczących przeszłych zachowań oraz jak można zwiększyć rzetelność otrzymywanych raportów?
- Czy procesy leżące u podstaw samoopisu zachowań i postaw zmieniają się w ciągu życia jednostki?
- Która technika jest najskuteczniejsza przy określaniu czy pytanie „działa” zgodnie z założeniami badawczymi?

Jak widać różnorodność pytań badawczych w tym obszarze jest ogromna. W prezentowanym opracowaniu skupimy się jedynie na czynniku przedmiotowym jaki stanowi wpływ różnych własności skal numerycznych na odpowiedzi udzielane przez uczestników badań kwestionariuszowych. Z uwagi na zainteresowania autorów opracowania problematyką jakości życia (Derbis, 2000; Derbis, Trawka, Wirga, 2004; Trawka, Derbis, 2006 w druku; Trawka, 2005; Wirga 2005) zajmiemy się w pierwszym rzędzie zmiennością efektów pomiaru poczucia szczęścia.

Celem przeprowadzonych przez nas badań jest analiza wpływu różnych właściwości skal numerycznych na zakres odpowiedzi udzielanych przez osoby badane. W szczególności interesuje nas odpowiedź na pytania:

- Jakie różnice będą w rozkładach uzyskanych odpowiedziach w zależności od tego, czy skala towarzysząca pytaniu będzie zaczynać się od liczby ujemnej, od zera lub od jedynki?
- Jakie przesunięcie w rozkładzie wyników nastąpi wówczas, gdy osoba badana ma pewną dowolność w interpretacji skali zaopatrzonej w niepełne etykiety opisujące wartości liczbowe skali?
- Jak zmienią się rozkłady odpowiedzi osób badanych, gdy „wydłużą się” skala możliwych odpowiedzi?

Przewidujemy, że przy zastosowaniu skali rozpoczynającej się od liczby ujemnej, odpowiedzi osób badanych będą się koncentrować bardziej „lewostronnie”, niż przy zastosowaniu skal rozpoczynających się od „zera” i „jedynki”. Przesunięcia koncentracji wyników na skali nie przesądzą o skośności potencjalnych rozkładów dlatego pojęcia skośności nie używamy.

Argumenty za takim rozkładem wyników przedstawiliśmy wyżej. Nadto wprowadzenie wartości ujemnych aktywizuje jakby szerszą przestrzeń fizyczną, którą odpowiadając można „zagospodarować” psychicznie, uwzględniając możliwości, jakie

daje dwubiegunowość. Fizyczny, neutralny środek całej skali, jednakowo oddalony od jej brzegów jest jednocześnie prawym krańcem wartości ujemnych a lewym wartości dodatnich, co sugeruje wymiar o przeciwnych cechach. Brak liczb ujemnych sugeruje konieczność oceny pozytywnego natężenia cechy. Ponadto, jeżeli osobom badanym nie narzucimy opisu etykiet każdego punktu na skali, ich odpowiedzi będą bardzo zróżnicowane i zależne od wartości (0, +, -) rozpoczynającej skalę numeryczną. Badany aktywizuje wtedy odpowiednią do wymiaru pamięć własnych opisów, zgodnych z jego kategoryzacją świata, która może zależeć od wymienionych czynników podmiotowych i przedmiotowych bardziej niż w sytuacji ustrukturalizowania punktów na skali etykietami. Ten efekt powinien więc być silniejszy w przypadku skali „wydłużonej”, niż w przypadku krótkich skal. Przez „wydłużenie” skali rozumiemy tutaj zastosowanie w jednych warunkach badawczych skali jedenastostopniowej (od -5 do 5), a w innych krótkiej, pięciostopniowej (od 1 do 5 lub od -2 do 2).

METODA

Aby odpowiedzieć na wymienione powyżej pytania, przeprowadziliśmy dwa badania. Badanie pierwsze miało na celu pomiar różnic w udzielanych odpowiedziach przez respondentów w zależności od wartości liczbowej rozpoczynającej skalę oraz od rodzaju etykiet opisujących wartości skali. Wykorzystano autorski kwestionariusz do pomiaru poczucia szczęścia. Jego charakterystyki psychometryczne nie są tu brane pod uwagę więc nie będą przedstawiane.

Badanie drugie dotyczyło zależności wyniku pomiaru od długości skali. Stanowi także częściową replikację wyników uzyskanych w badaniu pierwszym w tym sensie, że przeprowadzane było za pomocą innego, powszechnie znanego narzędzia badawczego (NEO-FFI). Przez porównanie wyników uzyskanych w obydwóch badaniach można było sprawdzić niezależność efektów pomiarów wynikających z właściwości skali od przedmiotu pomiaru.

BADANIE 1

Osoby badane. Przebadano dobranych losowo 180 studentów (113 kobiet i 67 mężczyzn; *M* wieku = 21.1 lat) wyższych uczelni znajdujących się na terenie miasta Opola (Uniwersytetu Opolskiego, Politechniki Opolskiej i Wyższej Szkoły Zarządzania i Administracji).

Materiały. Do badania został wykorzystany autorski kwestionariusz mierzący poczucie szczęścia skonstruowany na potrzeby innych badań dotyczących poczucia jakości życia. Narzędzie to składa się z 20 stwierdzeń dotyczących różnych aspektów życia jednostki i jej przeżyć. Odpowiedzi udzielane są na 5 stopniowej skali od zdecydowanie nie zgadzam się do zdecydowanie zgadzam się. Twierdzenia w kwestionariuszu nie ulegały zmianie – manipulowano natomiast wartościami skal numerycznych, które odpowiednio do danej grupy badawczej przyjmowały wartości:

Skala „ujemna”	-2	-1	0	1	2
----------------	----	----	---	---	---

lub

Skala „od zera”	0	1	2	3	4
-----------------	---	---	---	---	---

lub

Skala „od jedynki”	1	2	3	4	5
--------------------	---	---	---	---	---

Manipulowano także etykietami opisującymi wartości liczbowe, które odpowiednio do danej grupy badawczej przyjmowały postać:

Etykiety „pełne”

Zdecydowanie nie zgadzam się	Nie zgadzam się	Nie mam zdania	Zgadzam się	Zdecydowanie zgadzam się
I	II	III	IV	V

lub

Etykiety „pełne nad twierdzeniami”

- I – Zdecydowanie nie zgadzam się
- II – Nie zgadzam się
- III – Nie mam zdania
- IV – Zgadzam się
- V – Zdecydowanie się zgadzam

lub

Etykiety „na końcach skali”

Zdecydowanie nie zgadzam się				Zdecydowanie zgadzam się
I	II	III	IV	V

Rzymskie cyfry zastępowano odpowiednimi wartościami liczbowymi w zależności od grupy badawczej.

Przebieg badania. Osoby badane były przydzielone losowo do grup badawczych, odpowiadających brany przez nas pod uwagę właściwościom skal. Następnie zostały rozsadzone do oddzielnych stolików, aby siedziały pojedynczo i nie przeszkadzały sobie nawzajem. Przedstawiona została im wtedy informacja, że biorą udział w badaniach dotyczących poczucia szczęścia oraz że badania te są częścią większego projektu realizowanego przez Uniwersytet Opolski. Następnie badani dostali do anonimowego wypełnienia kwestionariusze poczucia szczęścia odpowiednie dla każdej z grup badawczych, tzn. zróżnicowane cechami skali wg przedstawionego wyżej schematu.

Wskaźnik zmiennej zależnej. Wprowadzono trzy poziomy pomiaru poczucia szczęścia. Liczbę odpowiedzi dotyczących I i II punktu skali (I – „Zdecydowanie nie zgadzam się” oraz II – „Nie zgadzam się”) przypisano poziomowi „1”. Liczbę odpowiedzi z III punktu skali (III – „Nie mam zdania”) przypisano poziomowi „2”. Odpowiedzi z IV i V punktu skali (IV – „Zgadzam się” oraz V – „Zdecydowanie zgadzam się”) przypisano poziomowi „3”. Zliczano na danym poziomie liczbę zakreśleń w całym 20-pytaniowym kwestionariuszu. Wskaźnik zmiennej zależnej stanowiła średnia liczba odpowiedzi na poszczególnych poziomach. Liczba ta przedstawiona jest także jako procent dokonanych zakreśleń. Przykładowo osoby badane na 20 możliwych zakreśleń w wersji z etykietami pełnymi i skalą od liczby „1”, zaznaczyły średnio „1,4” odpowiedzi na poziomie 1, co stanowi 13% odpowiedzi w kwestionariuszu (por. Tabela 1).

Tabela 1
Wyniki uzyskane w badaniu 1 (N = 180)

Zmienna niezależna: RODZAJ SKALI	Poziomy zmiennej zależnej	ETYKIETY PEŁNE			ETYKIETY PEŁNE NAD TWIERDZENIAMI			ETYKIETY NA KOŃCACH SKALI		
		M	SD	% odp.	M	SD	% odp.	M	SD	% odp.
SKALA	1	1.40	.94	13	1.20	1.01	12	1.20	1.36	12
OD	2	2.30	1.30	23	2.50	1.73	25	2.70	1.22	27
JEDYNKI	3	6.20	1.28	64	6.30	1.84	63	6.10	2.02	61
SKALA	1	1.10	1.33	11	1.70	1.66	17	1.30	1.89	13
OD	2	1.80	1.20	18	1.90	1.55	19	2.10	1.07	21
ZERA	3	7.10	1.55	71	6.40	2.11	64	6.60	2.39	66
SKALA	1	2.30	1.38	23	2.20	2.46	23	3.10	2.40	31
UJEMNA	2	1.00	1.12	10	1.40	1.05	14	2.00	1.12	20
	3	6.70	2.05	67	6.30	2.30	63	4.90	2.12	49

WYNIKI OTRZYMANE W BADANIU 1

Przeprowadzona analiza wg modelu ANOVA 3 (Skala: od wartości ujemnej vs od zera vs od jedynki) × 3 (Etykiety: pełne vs pełne nad twierdzeniami vs etykiety na końcach skali) pokazała, że porównywane grupy różnią się statystycznie istotnie [$F(12, 447) = 1.77; p < .05$]. Wyniki analizy kontrastów pomiędzy Skalami (zaczynające się od jedynki, zera, wartości ujemnej) dla poszczególnych Etykiet (pełne, pełne nad twierdzeniami, na końcach skali) są następujące:

- Dla „Etykiet pełnych” wykryto istotne statystycznie różnice w porównaniach wyników pomiaru: skala „od jedynki” × skala „ujemna” $F(3, 169) = 5,59; p < .001$ oraz skal „od zera” × skala „ujemna” $F(3, 169) = 2,65; p < .05$.
- Dla „Etykiet pełnych nad twierdzeniami” wykryto istotne różnice w porównaniach: skala „od jedynki” × skala „ujemna” $F(3, 169) = 4,84; p < .01$.
- Dla „Etykiet na końcach skali” wykryto istotne różnice w porównaniach: skala „od jedynki” × skala „ujemna” $F(3, 169) = 4,78; p < .01$ oraz skal „od zera” × skala „ujemna” $F(3, 169) = 3,78; p < .01$.

Wyniki analizy kontrastów pomiędzy Etykietami dla poszczególnych Skal:

- Dla skali „ujemnej” wykryto istotne statystycznie różnice w porównaniach: „Etykiety pełne” × „Etykiety na końcach skali” $F(3, 169) = 3,08; p < .05$ oraz „Etykiety pełne nad twierdzeniami” × „Etykiety na końcach skali” $F(3, 169) = 3,14; p < .05$.
- Dla skali „od zera” i skali „od jedynki” uzyskane różnice w wynikach pomiaru okazały

się nieistotne statystycznie bez względu na stosowaną Etykietkę.

Rozkład wszystkich wyników dla poszczególnych Skal i Etykiet przedstawia Tabela 1.

Wyniki analizy kontrastów dla poszczególnych podgrup, które różnią się statystycznie prezentują się następująco (por. Tabela 2):

- Poziom zmiennej zależnej 1, etykiety pełne, skala od zera vs Poziom zmiennej zależnej 1, etykiety pełne, skala ujemna ($p < .05$)
- Poziom zmiennej zależnej 2, etykiety pełne, skala od jedynki vs Poziom zmiennej zależnej 2, etykiety pełne, skala ujemna ($p < .01$)
- Poziom zmiennej zależnej 2, etykiety pełne, skala od zera vs Poziom zmiennej zależnej 2, etykiety pełne, skala ujemna ($p < .05$)
- Poziom zmiennej zależnej 2, etykiety nad twierdzeniami, skala od jedynki vs Poziom zmiennej zależnej 2, etykiety nad twierdzeniami, skala ujemna ($p < .01$)
- Poziom zmiennej zależnej 1, etykiety na końcach skali, skala od jedynki vs Poziom zmiennej zależnej 1, etykiety na końcach skali, skala ujemna ($p < .001$)
- Poziom zmiennej zależnej 1, etykiety na końcach skali, skala od zera vs Poziom zmiennej zależnej 1, etykiety na końcach skali, skala ujemna ($p < .001$)
- Poziom zmiennej zależnej 3, etykiety na końcach skali, skala od zera vs Poziom zmiennej zależnej 3, etykiety na końcach skali, skala ujemna ($p < .01$)
- Poziom zmiennej zależnej 2, etykiety pełne, skala ujemna vs Poziom zmiennej zależnej 2, etykiety na końcach skali, skala ujemna ($p < .05$)
- Poziom zmiennej zależnej 3, etykiety pełne, skala ujemna vs Poziom zmiennej zależnej 3, etykiety na końcach skali, skala ujemna ($p < .01$)

Tabela 2

Poziomy istotności dla poszczególnych grup badawczych

Poziomy zmiennej zależnej	Rodzaj skali	Rodzaj etykiety	ETYKIETY PEŁNE			ETYKIETY PEŁNE NAD TWIERDZENIAM			ETYKIETY NA KOŃCACH SKALI		
			Od 1	Od 0	Od -2	Od 1	Od 0	Od -2	Od 1	Od 0	Od -2
1	Od jedynki	Pełne	---			---					
		Nad twierdz.	ns			ns					
		Na końcach	ns			ns			---		
	Od zera	Pełne	ns	---		ns			ns		
		Nad twierdz.	ns	ns		ns	---		ns		
		Na końcach	ns	ns		ns	ns		ns	---	
Ujemna	Pełne	ns	.05	---	ns	ns		ns	ns		
	Nad twierdz.	ns	ns	ns	ns	ns	---	ns	ns		
	Na końcach	ns	ns	ns	ns	ns	ns	.001	.001	---	
2	Od jedynki	Pełne	---			---					
		Nad twierdz.	ns			---					
		Na końcach	ns			ns			---		
	Od zera	Pełne	ns	---		ns			ns		
		Nad twierdz.	ns	ns		ns	---		ns		
		Na końcach	ns	ns		ns	ns		ns	---	
Ujemna	Pełne	.01	.05	---	ns	ns		ns	ns		
	Nad twierdz.	ns	ns	ns	.01	ns	---	ns	ns		
	Na końcach	ns	ns	.05	ns	ns	ns	ns	ns	---	
3	Od jedynki	Pełne	---			---					
		Nad twierdz.	ns			---					
		Na końcach	ns			ns			---		
	Od zera	Pełne	ns	---		ns			ns		
		Nad twierdz.	ns	ns		ns	---		ns		
		Na końcach	ns	ns		ns	ns		ns	---	
Ujemna	Pełne	ns	ns	---	ns	ns		ns	ns		
	Nad twierdz.	ns	ns	ns	ns	ns	---	ns	ns		
	Na końcach	ns	ns	.01	ns	ns	.05	ns	.01	---	

- Poziom zmiennej zależnej 3, etykiety nad twierdzeniami, skala ujemna vs Poziom zmiennej zależnej 3, etykiety na końcach skali, skala ujemna ($p < .05$)
- Pozostałe interakcje okazały się nieistotne.

DYSKUSJA WYNIKÓW Z BADANIA 1

Otrzymane wyniki sugerują, że obydwa brane przez nas pod uwagę czynniki, tj. rodzaj etykiety i zastosowane wartości liczbowe opisujące skalę, wpływają na zróżnicowanie danych otrzymywanych w badaniach kwestionariuszowych. Czynniki te niekiedy wchodzi z sobą w interakcję. W niektórych przypadkach mogą także znosić swoje wzajemne oddziaływanie.

Zastosowanie skal „od jedynki” oraz „od zera” nie wpływa na otrzymane wyniki, bez względu na towarzyszący im rodzaj etykiet. Wywołują one bardzo zbliżony do siebie rozkład odpowiedzi udzielanych przez badanych. Wyniki uzyskane zarówno za pomocą skali rozpoczynającej się od zera jak i od jedynki potwierdzają wcześniejsze przypuszczenia badaczy, iż skale „dodatnie” komunikują badanemu jednowymiarowość analizowanych cech (por. Schwarz, 1991). Rozpoczynanie liczbowego opisu skali pięciostopniowej od „0” i od „1” nie różnicuje uzyskiwanych wyników.

Inne efekty powstają w przypadku zastosowania skali „ujemnej”. Jeżeli porównamy rozkłady odpowiedzi przypadających na pierwszy poziom zmiennej zależnej przy skalach „dodatnich” i skali „ujemnej” zauważymy znaczne przesunięcie wyników. Dla skali „dodatniej” zaczynającej się od „1” na pierwsze dwa punkty skali (poziom 1) przypada średnio 1,4 (13%) odpowiedzi, podczas gdy dla skali „ujemnej” jest to już 2,3 (25,6%). Wynik ten oznacza uświadomienie sobie przez badanego dwuwymiarowości skali ujemnej. Konsekwencją tego może być wzrost dostępności kategorii negatywnych w umyśle badanego i tym samym „poszerzenie” liczby kategorii odpowiedzi branych pod uwagę.

Z diagnostycznego punktu widzenia uzyskany efekt oznacza, że zastosowanie skali „ujemnej” pociąga za sobą obniżenie bezwzględnej wartości wyniku. Inaczej mówiąc wzrost liczby odpowiedzi po lewej stronie skali. W naszym przypadku oznaczałoby to diagnozę słabszego poczucia szczęścia osób badanych kwestionariuszem ze skalą ujemną, niż w przypadku zastosowania skal dodatnich.

W odniesieniu do częstości odpowiedzi „Nie mam zdania” dla skal „dodatnich” jest to średnio 2,13 (22,2%) odpowiedzi, podczas gdy dla skali „ujemnej” jest to tylko 1,2 (14,6%). Można zatem wnioskować,

że skala „ujemna” jako dwubiegunowa, wywołuje krystalizację opinii osób badanych o sobie.

Podobnie jak Schwarz, Knäuper, Hippler, Noelle-Neumann (1991) najbardziej znaczące różnice efektów pomiaru na różnych skalach zaobserwowaliśmy w przypadku „Etykiet na końcu skali”. Porównując skale „dodatnie” z „ujemną” w odniesieniu do zarówno skrajnych odpowiedzi potwierdzających (poziom „3”) i jak i przeczących (poziom „1”) obserwujemy istotne przesunięcie wyników. Dla skal „dodatnich” na pierwsze dwa punkty skali przypada średnio o 2/3 mniej odpowiedzi niż dla skali „ujemnej”. Z kolei na dwa ostatnie punkty skali, w przypadku skal „dodatnich”, przypada średnio o 1/4 odpowiedzi więcej niż dla tych samych odpowiedzi, tyle że umieszczonych na skali „ujemnej”. Różnice te można tłumaczyć wymiennymi wyżej właściwościami skali „ujemnej”, które są wzmacniane przez dowolność interpretacji brakujących etykiet opisujących poszczególne punkty skali. Podobne efekty, ale słabsze, wystąpiły także w przypadku dwóch pozostałych rodzajów etykiet, które mimo wszystko nie dają możliwości dowolnej interpretacji punktów skali osobom badanym.

Na podstawie przeprowadzonych badań, można by stwierdzić, że badana populacja jest mniej zadowolona ze swojego życia w przypadku skali „ujemnej” oraz „Etykiet na końcach skali” niż w przypadku skal „dodatnich” i „etykiet pełnych”. Efekt ten wskazuje na potencjalne źródło artefaktów. Dlatego przy wyciąganiu wniosków z wyników badań opierających się na samoopisie, a zwłaszcza przy ich porównywaniu, należy zwrócić szczególną uwagę na właściwości formalne zastosowanych skal – wszak mogą być źródłem czynników zakłócających rekonstrukcję rzeczywistości społecznej. Różnice wyników implikowane właściwościami skal nie wskazują jednak, jakie właściwości skal sprzyjają wierniejszym rekonstrukcją. To problem na osobne badania.

Wyjaśnienia przyczyn uzyskanych przez nas różnic wyników pomiaru mogą być także relacje pomiędzy zastosowanymi Etykietami a Skalą. Mogą one wzbudzać swoisty dysonans. Etykiety, jeśli towarzyszą każdej skali, bez względu na jej wartości liczbowe, sugerują, że analizowana cecha ma dwa bieguny: lewa strona (np.: nie zgadzam się) jest zaprzeczeniem prawej (np.: zgadzam się). Jednak skale liczbowe, które towarzyszą tym etykietom mogą sugerować badanemu co innego. W przypadku ujemnych wartości liczbowych i odpowiadających im etykiet z zaprzeczeniem (np. nie zgadzam się) oraz, co oczywiste, w przypadku liczbowych wartości dodatnich i odpowiadających im etykiet

potwierdzających (np. zgadzam się) dysonansu nie ma. Jednakże dodatnie wartości liczbowe na początku skali i odpowiadające im zaprzeczenia (1 – zdecydowanie się nie zgadzam) mogą powodować trudność. U badanego wzbudzone są bowiem negatywne treści semantyczne, którym odpowiadają dodatnie liczby. Mówiąc inaczej, sytuacje dwubiegunowych etykiet i jednobiegunowych skal dodatnich mogą powodować trudności w „umiejscowieniu się badanego na skali”, a w konsekwencji wzrost wyborów odpowiedzi „nie mam zdania”. W sytuacji, gdy występuje zgodność między skalą a etykietami (ujemna skala dwubiegunowa i dwubiegunowe etykiety) badani nie muszą uciekać się do neutralnej kategorii „nie mam zdania” i udzielają odpowiedzi bardziej zgodnych ze swoimi ocenami.

Można także zastanowić się, czy otrzymane wyniki nie są cechą charakterystyczną dla zastosowanego narzędzia, bądź czy nie mogą być wynikiem specyfiki pytań dotyczących poczucia jakości życia? Aby odpowiedzieć na to pytanie przeprowadziliśmy badanie 2, w którym wykorzystaliśmy kwestionariusz NEO-FFI Costy i McCrae w adaptacji Szczepaniaka, Śliwińskiej, Strelaua i Zawadzkiego (1998).

BADANIE 2 – METODA

Osoby badane. Przebadano 80 losowo dobranych słuchaczy studiów zaocznych (62 kobiety i 18 mężczyzn; średnia wieku $M = 27.2$ lat) kierunku pedagogika w Uniwersytecie Opolskim i Wyższej Szkole Zarządzania i Administracji w Opolu.

Materiały. Do badania wykorzystano trzy wersje kwestionariusza NEO-FFI – w wersji oryginalnej, w wersji ze skalą „ujemną” oraz w wersji ze skalą „wydłużoną” i „niepełnymi” etykietami. Arkusze testowe różniły się między sobą wartościami skali numerycznej odpowiedzi. W wersji niezmienionej skala możliwych odpowiedzi zaczyna się od 1 (zdecydowanie nie zgadzam się) do 5 (zdecydowanie zgadzam się). W wersji „ujemnej” skala możliwych odpowiedzi zaczyna się od –2 (zdecydowanie nie zgadzam się) do 2 (zdecydowanie zgadzam się). W wersji „wydłużonej” skala odpowiedzi rozpoczynała się od –5 (zdecydowanie nie zgadzam się) do 5 (zdecydowanie zgadzam się). Dodatkowo do tej skali dodane były tylko trzy etykiety opisujące wartości numeryczne. Etykieta – „zdecydowanie nie zgadzam się” odpowiadała wartości –5, „nie mam zdania” wartości 0 oraz „zdecydowanie zgadzam się” wartości 5.

Przebieg badania. Badani zostali rozsadzeni do osobnych stolików i rzędami przydzieleni do jednej z dwóch grup badawczych. Następnie otrzymali do anonimowego wypełnienia kwestionariusz NEO-FFI. Grupa 1 wypełniała kwestionariusz w wersji „ujemnej”. Grupa 2 wypełniała niezmieniony kwestionariusz. Grupa 3 wypełniała kwestionariusz w wersji „wydłużonej”. Po wypełnieniu testów prowadzący zbierali testy, dziękowali i wychodzili z sali.

Wskaźnik zmiennej zależnej. Wskaźnikiem zmiennej zależnej, rozumianej jako rozkład odpowiedzi na skalach kwestionariusza NEO-FFI, była liczba odpowiedzi przypadająca na poszczególne wartości skali. W przypadku kwestionariusza w wersji standardowej i ze skalą „ujemną” pomiaru zmiennych zależnych dokonano w ten sam sposób jak w badaniu 1. Natomiast w przypadku kwestionariusza w wersji „wydłużonej” odpowiedzi z punktu skali od –5 do –1 przypisano poziomowi „1”. Odpowiedzi z punktu 0 na skali (0 – „Nie mam zdania”) przypisano poziomowi „2”. Odpowiedzi z punktów skali od 1 do 5 przypisano poziomowi „3”.

Dodatkowo przeliczano wyniki surowe uzyskane w kwestionariuszu na wyniki stenowe zgodnie z kluczem do kwestionariusza NEO-FFI. Wyniki uzyskane w kwestionariuszu ze skalą „ujemną” potraktowano tak, jakby pochodziły ze skali niezmienionej. Przeliczono je analogicznie jak kwestionariusz oryginalny. W przypadku kwestionariusza ze skalą „wydłużoną”, zastosowano następującą procedurę: wartościom „–5,–4” odpowiadała wartość „1” w oryginalnym kwestionariuszu NEO-FFI, wartościom „–3,–2,–1” odpowiadała wartość „2”, wartości „0” wartość „3”, wartościom „1,2,3” wartość „4” i ostatecznie wartościom „4” i „5” odpowiadała wartość 5 w kwestionariuszu NEO-FFI. Trzeba zaznaczyć, że przeliczanie w ten sposób wartości nie służyło rzeczywistemu diagnozowaniu osób badanych, a jedynie zaprezentowaniu różnic rozkładu odpowiedzi w przypadku zmiany cech skali.

WYNIKI OTRZYMANE W BADANIU 2

Analiza statystyczna wykazała istotne różnice pomiędzy analizowanymi wersjami kwestionariusza NEO-FFI [$F(6, 148) = 14.511, p < .0001$]. Rozkład wyników uzyskanych w badaniu drugim prezentuje Tabela 3. Analiza kontrastów dla oczekiwanych średnich brzegowych wykazała następujące zależności:

- Różnice pomiędzy wynikami pomiaru skalą standardową a skalą „ujemną” z zachowaną

długością są istotne statystycznie [$F(3, 74) = 3.921; p < .05$]

- Różnice pomiędzy wynikami pomiaru skalą standardową a „ujemną skalą wydłużoną” są istotne statystycznie [$F(3, 74) = 35.137; p < .0001$]
- Różnice pomiędzy wynikami pomiaru skalą „ujemną” a „ujemną skalą wydłużoną” są istotne statystycznie [$F(3, 74) = 15.933; p < .0001$]

Wyniki analizy kontrastów dla poszczególnych podgrup prezentują się następująco (por. Tabela 4):

- Poziom zmiennej zależnej 1, skala standardowa vs poziom zmiennej zależnej 1, skala „ujemna wydłużona” ($p < .05$)
- Poziom zmiennej zależnej 2, skala standardowa vs poziom zmiennej zależnej 2, skala „ujemna 5-punktowa” ($p < .01$)
- Poziom zmiennej zależnej 2, skala „ujemna 5-punktowa” vs poziom zmiennej zależnej 2, skala „ujemna wydłużona” ($p < .001$)
- Poziom zmiennej zależnej 2, skala standardowa vs poziom zmiennej zależnej 2, skala „ujemna wydłużona” ($p < .001$)

- Poziom zmiennej zależnej 3, skala „ujemna 5-punktowa” vs poziom zmiennej zależnej 3, skala „ujemna wydłużona” ($p < .001$)
- Poziom zmiennej zależnej 3, skala standardowa vs poziom zmiennej zależnej 3, skala „ujemna wydłużona” ($p < .001$)
- Pozostałe interakcje okazały się nieistotne.

Przeprowadzono także dodatkową analizę sprawdzającą różnice między analizowanymi grupami pod względem każdego z pięciu czynników osobowości wg Costy i McCrae w zależności od zastosowanej wersji kwestionariusza. Uzyskane wyniki wykazały różnice istotne statystycznie [$F(10, 144) = 4.649; p < .0001$]. Analizy post-hoc (test Tukeya) dla każdego czynnika osobowości wykazały następujące zależności.

Dla Neurotyzmu:

- Pomiedzy wynikami uzyskanymi za pomocą kwestionariusza w wersji standardowej a wynikami pomiaru kwestionariuszem w wersji „ujemnej wydłużonej” istnieją istotne statystycznie różnice ($p < .01$)

Tabela 3

Rozkład wyników uzyskanych w badaniu 2

	POZIOM ZMIENNEJ ZALEŻNEJ								
	1			2			3		
	M	SD	% odp.	M	SD	% odp.	M	SD	% odp.
Skala standardowa (od 1 do 5)	16.93	3.65	28	15.10	6.98	25	27.96	6.13	47
Skala ujemna (od -2 do 2)	18.80	4.28	32	10.36	3.81	17	30.84	4.48	51
Skala wydłużona (od -5 do 5)	20.17	4.24	34	1.87	2.86	3	37.54	4.04	63

Tabela 4

Poziomy istotności dla poszczególnych grup badawczych

Rodzaj skali	Poziom zmiennej zależnej	1			2			3		
		Od 1	Od -2	Od -5	Od 1	Od -2	Od -5	Od 1	Od -2	Od -5
Skala standardowa (od 1 do 5)	1	---								
	2				---					
	3							---		
Skala ujemna (od -2 do 2)	1									
	2		---		.01		---			
	3								---	
Skala wydłużona (od -5 do 5)	1	.05		---						
	2				.001	.001	---			
	3							.001	.001	---

- Pomiędzy wynikami kwestionariusza w wersji „ujemnej 5-punktowej” i wersji „ujemnej wydłużonej” są istotne statystycznie różnice ($p < .05$)
- Pomiędzy wynikami zebranych kwestionariuszem w wersji oryginalnej i wersji „ujemnej 5-punktowej” nie odnotowano istotnych różnic.

Dla Ekstrawersji:

- Nie ma istotnych różnic pomiędzy uzyskanymi wynikami bez względu na zastosowaną wersję kwestionariusza.

Dla Otwartości:

- Pomiędzy wynikami kwestionariusza ze skalą „ujemną 5-punktową” i wynikami kwestionariusza ze skalą „ujemną wydłużoną” są istotne różnice ($p < .001$), podobnie jest w przypadku skali standardowej i skali „ujemnej wydłużonej” ($p < .01$)

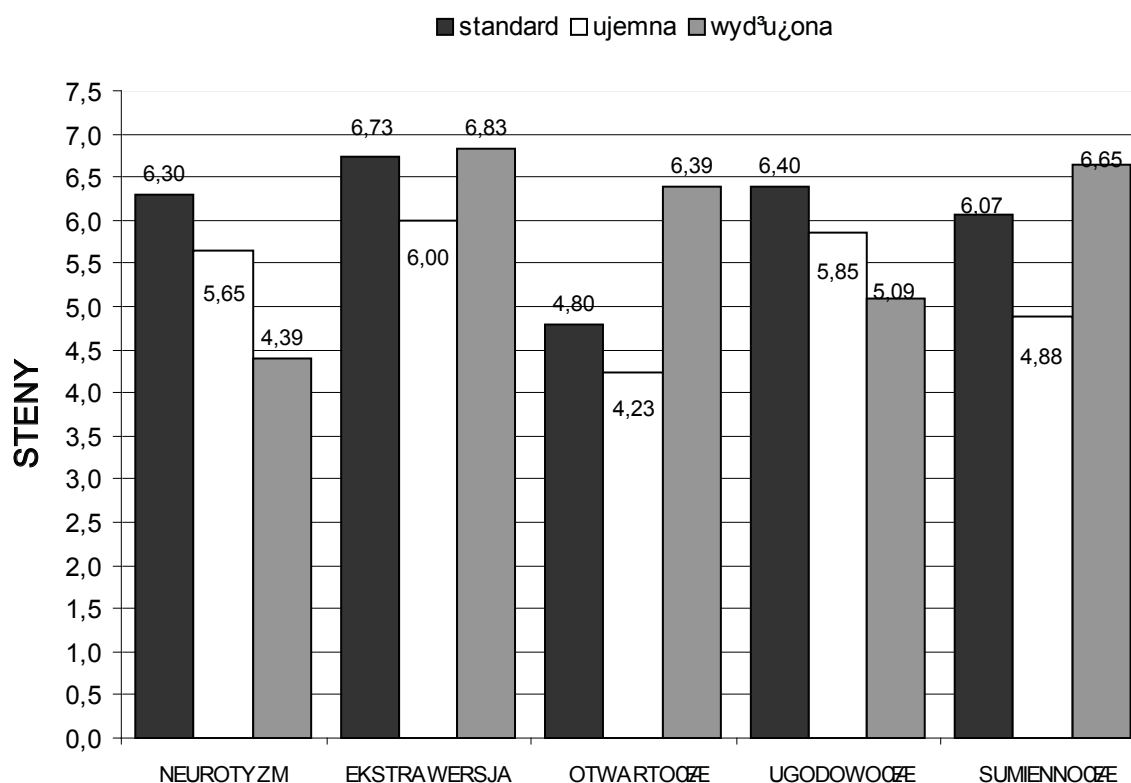
- Nie ma różnic w wynikach uzyskanych za pomocą kwestionariusza w wersji standardowej a wersją „ujemną 5-punktową”

Dla Ugodowości:

- Pomiędzy wynikami kwestionariusza w wersji standardowej a wynikami kwestionariusza w wersji „ujemnej wydłużonej” istnieją istotne statystycznie różnice ($p < .01$)
- Pozostałe wersje kwestionariusza nie różnicują istotnie wyników.

Dla Sumienności:

- Pomiędzy wynikami kwestionariusza ze skalą „ujemną 5-punktową” a wynikami kwestionariusza ze skalą oryginalną są istotne różnice ($p < .05$)
- Pomiędzy wynikami kwestionariusza ze skalą „ujemną 5-punktową” a wynikami kwestionariusza ze skalą „ujemną wydłużoną” są istotne różnice ($p < .01$)



Wykres 1. Różnice pomiaru pięciu czynników osobowości w zależności od zastosowanej wersji skali w teście NEO-FFI

- Nie ma różnic w wynikach uzyskanych za pomocą kwestionariusza w wersji standardowej a wersją „ujemną wydłużoną”

Wykres 1 obrazuje różnice pomiaru pięciu czynników osobowości w zależności od zastosowanej wersji kwestionariusza.

DYSKUSJA WYNIKÓW Z BADANIA 2

Ogólnie patrząc wyniki otrzymane w badaniu 2 potwierdziły zależności uzyskane w badaniu 1 – tzn. efekt pomiaru zmiennych zależy od formalnych właściwości zastosowanych skal.

Porównując rozkłady wyników uzyskanych za pomocą trzech branych pod uwagę wersji skal można zaobserwować dla wersji „ujemnej z zachowaną długością” i wersji „ujemnej wydłużonej” w stosunku do wersji standardowej efekt wzrostu średniej liczby odpowiedzi przypadających na poziom pierwszy („nie zgadzam się”) i trzeci („zgadzam się”) zmiennej zależnej. Natomiast średnia liczba odpowiedzi przypadających na poziom drugi („nie mam zdania”) zmiennej zależnej maleje. Przyczyn tego zjawiska może być kilka. Po pierwsze, indukcja rozległych reprezentacji poznawczych powodowana skalami ujemnymi sprzyja krystalizacji opinii na swój temat, dlatego osobom badanym łatwiej jest określić, czy zgadzają się, czy też nie zgadzają się z danym stwierdzeniem. Nie muszą wtedy sięgać do kategorii neutralnej. Po drugie, skale ujemne, w przeciwieństwie do oryginalnych skal dodatnich kwestionariusza NEO-FFI, są semantycznie zgodne z dwubiegunowymi etykietami dołączonymi do skali. Dzięki temu nie budzą dysonansu u osób badanych, przez co łatwiej im jest udzielić odpowiedzi wyrazistych (zgadzam się/nie zgadzam się). Po trzecie, na podstawie uzyskanych wyników można także powiedzieć, że wartości ujemne, ponieważ sugerują dwubiegunowość analizowanego wymiaru, to faworyzują pozytywne lub negatywne ustosunkowanie się badanego do poszczególnych stwierdzeń.

Jeżeli przeanalizujemy szczegółowe różnice pomiędzy wynikami uzyskanymi za pomocą skali „ujemnej” 5-punktowej i skali „ujemnej wydłużonej” 11-punktowej, to dojdziemy do wniosku, że im dłuższa jest skala ujemna, tym wyniki ulegają większemu „odsunięciu” od wartości środkowej – „nie mam zdania” w stronę prawą lub lewą na skali odpowiedzi. Sugerować to może, że wydłużenie skali odpowiedzi nie tylko fizycznie i wizualnie zwiększa liczbę możliwości, jakie pojawiają się przed badanym, ale także zwiększa dostępność i rozległość odmiennych kategorii semantycznych

u umysłach badanych. W rezultacie, badany nie tylko nie musi odwoływać się do odpowiedzi neutralnej, i tak naprawdę mało diagnostycznej, ale korzystając z uzyskanej swobody opisu siebie, wybiera świadomie odpowiedzi bardziej dla siebie informacyjne.

Inne wytłumaczenie uzyskanych wyników nasuwa analiza różnic w opisie wartości na każdej skali. Etykiety towarzyszące skali „ujemnej” 5-punktowej były przypisane do każdego punktu na skali a w przypadku skali „ujemnej wydłużonej” 11-punktowej etykiety znajdowały się jedynie na punktach krańcowych oraz przy wartości środkowej „nie mam zdania”. Uzyskany rozkład wyników w obrębie obu skal „ujemnych” może być przyczyną tego, że przy 11-punktowej skali odpowiedzi, badani sami tworzyli sobie etykiety dotyczące nie opisanych punktów skali, przez co przywoływane przez nich kategorie były bardzo indywidualne, co ułatwiało im wybranie odpowiedzi innej niż „nie mam zdania”. Trudność w interpretacji efektów pomiaru uzyskanych za pomocą „ujemnej skali wydłużonej” z niepełną liczbą etykiet jest niemożność ustalenia w tych badaniach w jakim stopniu uzyskane efekty zależą od długości skali ujemnej, a w jakim od stosowania niepełnych etykiet. Będzie to przedmiotem dalszych poszukiwań.

Zastosowany schemat naszych badań jest taki, że podstawowa część rezultatów łączy się z wprowadzeniem skal ujemnych. Pomiar przeprowadzone za pomocą skal w wersji oryginalnej i w wersjach ze skalami ujemnymi pokazują, że w przypadku zastosowania skali ujemnej z zachowaną długością (5-punktów) we wszystkich skalach wyniki ulegają obniżeniu, aczkolwiek istotne statystycznie różnice występują jedynie w przypadku skali „sumienność”. W przypadku skali „ujemnej wydłużonej” (11-punktów) uzyskane wyniki pomiaru dla poszczególnych skal są bardziej zróżnicowane. „Ujemna skala wydłużona” różnicuje wyniki, ale nie zawsze w tym samym kierunku. W odniesieniu do skali Neurotyzmu i Ugodowości obniża je, ale statystycznie istotnie tylko dla tej pierwszej. Na skali Otwartości i Sumienności wyniki statystycznie istotnie rosną a pomiar Ekstrawersji nie zależy od wprowadzonych warunków. Na wyniki stenowe w wersji „ujemnej wydłużonej” kwestionariusza NEO-FFI mógł wpłynąć sposób przeliczania wyników surowych na odpowiadające im stopnie w oryginalnej skali NEO-FFI. Możliwe, że inny sposób przeliczenia danych dałby dla poszczególnych czynników osobowości nieco inne rozkłady.

Należy zauważyć, że wartości cech osobowości przedstawione w stenach mieszczą się zazwyczaj w przedziale wartości przeciętnych. Jednak w za-

leżności od własności zastosowanej skali wyniki jednostkowe ulegają znacznemu przesunięciu.

Przeprowadzone porównania efektów pomiarów osobowości pokazują wprawdzie, że zależą one od formalnych właściwości użytych skal, ale nie wszystkie czynniki jednakowo mocno poddają się tej manipulacji. Dla niniejszych rozważań nie jest tak ważne, które są to czynniki, jak ważne jest zwrócenie uwagi, że na zmienność wyniku pomiaru pewien wpływ ma także to, co się mierzy (neurotyzm, ekstrawersję, sumienność...), nie tylko własności skali.

PODSUMOWANIE I WNIOSKI KOŃCOWE

W naukach społecznych stosowanie skal samoopisowych pomimo ich słabych stron to codzienność badawcza. Zazwyczaj badacze mają jakąś świadomość zależności efektów pomiaru od właściwości zastosowanych skal. Uzyskane przez nas wyniki mogą w skromnym stopniu podnieść świadomość metodologiczną badacza w tym względzie i jego wrażliwość na ten problem. Nie za dużo mówią o wartości diagnostycznej każdej z zastosowanych przez nas wersji. Nie wiadomo na ich podstawie, który rodzaj skali jest najlepszy, czyli który najlepiej rekonstruuje rzeczywistość społeczną, ponieważ nie o to chodziło w naszych badaniach. To problem do dalszych badań. Wiadomo jednak, że rodzaj skali i etykiet mogą różnicować rozkład odpowiedzi.

Zasadnicze wyniki dwóch naszych eksperymentów sugerują, że „skale ujemne” w stosunku do „skal dodatnich” sprzyjają wśród osób badanych tendencji do unikania opcji „nie mam zdania” na rzecz pozostałych możliwości odpowiadania, zwłaszcza na rzecz opcji po „lewej stronie” skali. Opcja „nie mam zdania”, z punktu widzenia badacza, ma małą wartość diagnostyczną i informacyjną, dlatego korzystniejsze jest z jego punktu widzenia, aby osoby badane wybierały bardziej wyraziste odpowiedzi.

W świetle uzyskanych wyników pozwalamy sobie na sugestię zachowywania zgodności pomiędzy wartościami numerycznymi skali a treścią etykiet. Aby nie wywołać dysonansu i ułatwić osobie badanej możliwie adekwatne uplasowanie się na skali, dwubiegunowe etykiety (od nie zgadzam się do zgadzam się) powinny być połączone z dwubiegunową skalą (od wartości ujemnych do dodatnich). W takich warunkach badawczych badany odwołuje się do kategorii semantycznych mniej lub bardziej nasyconych negatywnie lub mniej lub bardziej nasyconych pozytywnie, wspartych odpowiednio liczbowymi wartościami ujemnymi lub dodatnimi.

Uzyskiwane wyniki pomiaru mogą zmienić się także w zależności od rodzaju etykiet opisujących wartości na skali. O ile rodzaj etykiety (pełne, pełne nad twierdzeniami, na końcach skali) nie specjalnie różnicują wyniki pomiaru w przypadku „skal dodatnich” to różnicują efekty w przypadku „skal ujemnych”. I tak etykiety umieszczone na końcach „skali ujemnej” w stosunku do takiej samej długości „skal dodatnich” (zaczynających się tak od „0” jak i od „1”) sprzyjają większej częstości zakreśleń po lewej stronie skali, a mniejszej po prawej stronie. Z tego punktu widzenia powszechne stosowanie „skal dodatnich” wydaje się być o tyle bezpieczne, że wynik pomiaru dokonany za ich pomocą słabo zależy od cech formalnych skal. Jednakże problemy mogą powstać przy próbie replikacji cudzych badań. Badacze w opisie narzędzia zazwyczaj podają jaki zakres miała zastosowana przez nich skala, ale rzadko podają od jakiej wartości się zaczynała i jakie etykiety stosowano. Stąd jawi się postulat edytorski związany z oczekiwaniem dokładnego opisu właściwości formalnych stosowanych skal. Odmienność wyników replikowanych badań może mieć źródła w nieświadomym stosowaniu skal różniących się w tym względzie.

Świadomość powstawania różnic efektów pomiaru jako następstwa zastosowania różnych etykiet czy skal może także prowadzić do nieetycznych zachowań badaczy. Chodzi o wybieranie do swoich badań takich wartości skali czy takich etykiet, który zwiększą prawdopodobieństwo potwierdzenia hipotez badawczych. Miejmy nadzieje, że wzrost świadomości różnic płynących z zastosowania określonych cech formalnych kwestionariuszy będzie służyć podniesieniu rzetelności zbieranych danych, a nie będzie zaś służyć manipulowaniu wynikami.

LITERATURA CYTOWANA

- Aranowska, E. (1996). *Metodologiczne problemy zastosowań modeli statystycznych w psychologii*. Warszawa: Studio1.
- Brzeziński, J. (1978). *Metodologiczne i psychologiczne wyznaczniki procesu badawczego w psychologii*. Poznań: Wydawnictwo UAM.
- Brzeziński, J. (1978a). *Elementy metodologii badań psychologicznych*. Warszawa: Wydawnictwo naukowe PWN.
- Brzeziński, J. (1984). *Wybrane zagadnienia z psychometrii i diagnostyki psychologicznej*. Poznań: Wydawnictwo Naukowe UAM.
- Brzeziński, J. (red.). (1993). *Psychologiczne i psychometryczne problemy diagnostyki psychologicznej*. Poznań: Wydawnictwo Naukowe UAM.
- Brzeziński, J. (1993a). *Wymiary przestrzeni diagnostycznej*. W: J. Brzeziński (red.), *Psychologiczne i psychometryczne problemy diagnostyki psychologicznej* (s. 7–29). Poznań: Wydawnictwo Naukowe UAM.

- Brzeziński, J. (1996). *Metodologia badań psychologicznych*. Warszawa: PWN.
- Brzeziński, J. (2000). *Badania eksperymentalne w psychologii i pedagogice*. Warszawa: Wydawnictwo Naukowe „Scholar”.
- Brzeziński, J., Hornowska, E. (red.). (1993). *Z psychometrycznych problemów diagnostyki psychologicznej*. Poznań: Wydawnictwo Naukowe UAM.
- Brzeziński, J., Siuta, J. (red.). (2000). *Spółeczny kontekst badań psychologicznych i pedagogicznych. Wybór tekstów*. Poznań: Wydawnictwo Naukowe UAM.
- Brzeziński, J., Toeplitz-Winiewska, M. (red.). (2000). *Etyczne dylematy psychologii*. Poznań: Wydawnictwo Fundacji Humaniora.
- Churchill, G. (2002). *Badania marketingowe. Podstawy metodologii*. Warszawa: PWN.
- Czapiński, J. (1994). Uziemienie polskiej duszy. *Kultura i Społeczeństwo*, 3, 17–37.
- Czapiński, J. (1998). *Jakość życia Polaków w czasie zmiany społecznej. Raport końcowy z realizacji projektu badawczego*. Warszawa: Międzywydziałowy Instytut Studiów Społecznych UW.
- Derbis, R. (2000). *Doświadczenie codzienności. Poczucie jakości życia, swoboda działania, odpowiedzialność, wartości osób bezrobotnych*. Częstochowa: Wydawnictwo Wyższej Szkoły Pedagogicznej.
- Derbis, R., Trawka, K., Wirga, T. (2004). Wpływ etykiety i rodzaju skali numerycznej na uziemienie i lewitację duszy polskiej. *Prace naukowe Akademii im. Jana Długosza w Częstochowie. Psychologia XII* (s. 59–66). Częstochowa: Wydawnictwo Akademii im. J. Długosza.
- Drwal, R. (1989). *Techniki kwestionariuszowe w diagnostyce psychologicznej. Wybrane zagadnienia*. Lublin: Uniwersytet Marii Curie-Skłodowskiej.
- Drwal, R. (1995). *Adaptacja kwestionariuszy osobowości*. Warszawa: PWN.
- Greenleaf, E.A. (1992). Measuring extreme response style. *Public Opinion Quarterly*, 56, 328–351.
- Kordos, M. (1994). *Wykłady z historii matematyki*. Warszawa: WSiP.
- Krysicki, W. (1973). *Jak liczono dawniej, a jak liczymy dzisiaj*. Warszawa: Nasza Księgarnia.
- Nowakowska, M. (1975). *Psychologia ilościowa z elementami naukometrii*. Warszawa: PWN.
- Nowakowska, M. (1977). *Teorie badań. Ujęcie modelowe*. Warszawa: PWN.
- Paluchowski, Wł., J. (Red.). (1983). *Z zagadnień diagnostyki osobowości*. Wrocław: Zakład Narodowy Imienia Ossolińskich, Wydawnictwo PAN.
- Paluchowski, Wł., J. (1991). *Diagnozowanie osobowości. Testowanie-interpretacja-intervencja*. Poznań: Wydawnictwo Naukowe UAM.
- Schwarz, N. (1999). Self-Reports: How the questions shape the answers. *American Psychologist*, 54, 2, 93–105.
- Schwarz, N., Hippler, H.J. (1995). The numeric values of rating scales: A comparison of their impact in mail surveys and telephone interviews. *International Journal of Public Opinion Research*, 7, 72–74.
- Schwarz, N., Hippler, H.J., Deutsch, B., Starck, F. (1985). Response scales: Effects of category range on reported behavior and comparative judgments. *Public Opinion Quarterly*, 49, 388–395.
- Schwarz, N., Knäuper, B., Hippler, H.J., Noelle-Neumann, E. i Clark, F. (1991). Rating scales: Numeric values may change the meaning of scale labels. *Public Opinion Quarterly*, 55, 570–582.
- Trawka, K. (2005). Jak afekt i samoocena kształtują jakość życia – badania eksperymentalne. *Prace naukowe Akademii im. Jana Długosza w Częstochowie. Psychologia XIII*. Częstochowa: Wydawnictwo Akademii im. J. Długosza.
- Trawka, K., Derbis, R. (2006, w druku). Samoocena, afekt i nastrój jako wyznaczniki poczucia jakości życia. *Psychologia Jakości Życia*, 1.
- Wirga, T. (2005). Zmienne osobowościowe i temperamentalne jako wyznaczniki poczucia jakości życia. *Prace Naukowe Akademii im. Jana Długosza w Częstochowie. Psychologia XIII*. Częstochowa: Wydawnictwo Akademii im. J. Długosza.
- Zawadzki, Z., Strelau, J., Szczepaniak, P., Śliwińska, M. (1998). *Inwentarz osobowości NEO-FFI Costy i McCrae*. Warszawa: Pracownia Testów Psychologicznych PTP.