

# Bohdan Dudek

---

## Analiza testów psychologicznych zastosowanych do oceny respondentów w wywiadzie kwestionariuszowym

---

Przegląd Socjologiczny Sociological Review 30, 113-121

---

1978

Artykuł został zdigitalizowany i opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej [bazhum.muzhp.pl](http://bazhum.muzhp.pl), gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

BOHDAN DUDEK  
Uniwersytet Łódzki

## ANALIZA TESTÓW PSYCHOLOGICZNYCH ZASTOSOWANYCH DO OCENY RESPONDENTÓW W WYWIADZIE KWESTIONARIUSZOWYM

Treść: Wprowadzenie. — Analiza wyników. — Dyskusja i wnioski.

### WPROWADZENIE

Artykuł niniejszy jest kontynuacją artykułu B. Dudka i R. Kietlińskiego pt. *Zastosowanie testów psychologicznych do oceny osób odpowiadających na pytania w wywiadzie kwestionariuszowym*<sup>1</sup>. Autorzy przedstawili w nim jeden ze sposobów ustalenia stopnia zaufania do wyników otrzymanych za pomocą wywiadu kwestionariuszowego. Polega on na zmierzeniu pewnych cech respondenta, które mogą powodować udzielanie przez niego prawidłowych<sup>2</sup> bądź nieprawidłowych odpowiedzi. Jeżeli badacz ustali w sposób empiryczny, że jakaś cecha jest związana z udzielaniem prawidłowych odpowiedzi na dane pytanie, będzie mógł oszacować jakość uzyskanych wyników na podstawie stopnia, w jakim grupa jest nasycona daną cechą.

Przyjęto, że takimi cechami wpływającymi na proces udzielania odpowiedzi i jego wynik są: niezrozumienie lub złe zrozumienie pytania, tendencja do pokazywania się w lepszym świetle i skłonność do odpowiadania bez głębszego zastanawiania się nad treścią pytania. Do pomiaru tych właściwości przygotowano odpowiednie testy: test znajomości słów (S), test „kłamstwa” (K) oraz zestaw pytań „pustych” (P)<sup>3</sup>.

---

<sup>1</sup> Artykuł ten ukazał się w pracy zbiorowej *Analizy i próby technik badawczych w socjologii*, red. Z. Gostkowski, J. Lutyński, t. 5, Wrocław 1975.

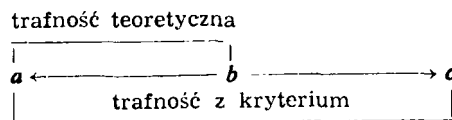
<sup>2</sup> Odpowiedź prawidłowa to odpowiedź stanowiąca rezultat procesów (psychicznych i innych) założonych przez badacza. Odpowiedź taka powinna być trafna, tj. powinna przynosić prawdziwą informację poszukiwaną przez badacza. Patrz J. Lutyński, *Analizy weryfikacyjne w badaniach z zastosowaniem wywiadu kwestionariuszowego, ich rodzaje i możliwości*, w: *Analizy i próby...*, t. 5.

<sup>3</sup> Pytanie „puste” to takie pytanie, które nie spełnia założeń poprawności

Analiza wyników wykazała, że rezultaty uzyskane przez badanych w testach skorelowane są dość silnie z ilością udzielanych przez nich prawidłowych odpowiedzi. 81% badanych, którzy otrzymali w teście *S* wyniki powyżej średniej, a w testach *K* i *P* wyniki poniżej średniej, udzieliło większą niż przeciętna ilość prawidłowych odpowiedzi. Natomiast 97% osób uzyskujących niskie wyniki w teście *S* oraz wysokie wyniki w testach *K* i *P* nie przekracza średniej ilości prawidłowych odpowiedzi. Można więc na podstawie wyników testowych przewidywać, którzy respondenci udzielać będą więcej, a którzy mniej prawidłowych odpowiedzi.

Artykuł jest próbą określenia trafności wyżej wymienionych testów na podstawie kryterium, na które składała się ilość prawidłowych odpowiedzi udzielanych przez badanych na pytania kwestionariuszowe.

Badanie trafności opiera się na następującym rozumowaniu: jeżeli istnieje związek o określonej sile pomiędzy wynikami testu a pewnymi właściwościami lub procesami wewnętrznymi osobnika oraz jeżeli istnieje związek o określonej sile pomiędzy właściwościami lub procesami wewnętrznymi osobnika a pewnym zewnętrznym zachowaniem (kryterium), to powinien istnieć związek pomiędzy wynikami testu a zewnętrznym zachowaniem. Siła tego ostatniego związku zależy od siły poprzednich związków. Rozumowanie to można ująć w schemat.



- a* — wyniki testowe,  
*b* — właściwości i procesy wewnętrzne,  
*c* — zachowanie, które stanowi kryterium dla *a* lub *b*,  
 strzałki oznaczają związek i kierunek wpływu.

Na schemacie zaznaczono dwa zasadnicze rodzaje trafności. Badacz, chcąc ustalić trafność testu na podstawie kryterium, określa w sposób empiryczny siłę związku pomiędzy wynikami testowymi (*a*) uzyskanymi przez badanych a ich zachowaniem w określonej sytuacji (*c*). Na powyższym schemacie można zauważyć wyraźnie, że trafność związana z kryterium opiera się na uproszczonym rozumowaniu. Jest to efekt behawiorystycznej tendencji do pomijania w badaniach zmiennych pośredniczących. Nie znaczy to jednak, że ustalenie trafności jest pozbawione sensu czy nieuprawnione. Jest to jeden z zasadniczych parametrów testu. Natomiast należy przyjąć, że trafność związana z kryterium nie mówi nam, co test mierzy, a tylko pozwala przewidywać określone

logicznej lub którego założenia empiryczne nie są spełnione. Na takie pytanie jedyną prawidłową odpowiedzią jest wskazanie błędu zawartego w pytaniu.

zachowania ludzi z pewnym prawdopodobieństwem. Pomiary kryterium dokonywane są bowiem innymi metodami. Związek pomiędzy wynikami testowymi a wynikami kryterialnymi jest upośredniony poprzez zmienne wewnętrzne (właściwości, procesy), które mają postać konstruktów hipotetycznych.

Badanie związku pomiędzy wynikami testowymi a własnościami lub procesami zachodzącymi wewnątrz jednostki, które mają charakter teoretyczny (hipotetyczny) nazywa się w psychometrii badaniem trafności teoretycznej testu<sup>4</sup> (na naszym schemacie jest to związek pomiędzy  $a$  i  $b$ ). Jednak badania tego nie należy ograniczać tylko do przypadków, w których nie możemy uzyskać zewnętrznego, obserwowalnego kryterium. Znajomość tego, z jakimi procesami wewnętrznymi korelują wyniki testu, jest ważna, jeżeli badacz chce włączyć interpretację wyników testowych w ramy swojej teorii.

Pozostał jeszcze do wyjaśnienia związek pomiędzy procesami wewnętrznymi ( $b$ ) a obserwowalnym kryterium ( $c$ ). Można powiedzieć o nim tylko tyle, że w konkretnym badaniu jest on przyjmowany na podstawie dotychczasowej wiedzy badacza albo wynika z definicji jednego z członów tego związku. Możemy bowiem definiować proces wewnętrzny ( $b$ ) jako taki proces, który prowadzi do określonego zachowania ( $c$ ) albo możemy definiować kryterium ( $c$ ) jako skutek zaistnienia poprzednio procesu wewnętrznego ( $b$ ). Na przykład w naszych badaniach związek pomiędzy tendencją do pokazywania się w lepszym świetle (tj.  $b$ ) a udzielaniem prawidłowej odpowiedzi (tj.  $c$ ) wynika z jej definicji, która mówi, że odpowiedź prawidłowa pojawia się w rezultacie założonych przez badacza procesów zachodzących u respondenta i dlatego przynosi prawdziwą informację. Tak więc odpowiedź prawidłową otrzymamy wtedy, gdy badany nie będzie odpowiadał tak, aby lepiej wypaść w oczach ankietera<sup>5</sup>.

Z przedstawionego schematu wynika, że w poprzednim artykule określono trafność zastosowanych testów na podstawie kryterium. Wykazano, że wyniki w testach: słownikowy, „kłamstwa” i pytań „pustych”, związane są z ilością prawidłowych odpowiedzi udzielanych przez badanych, czyli określono siłę związku pomiędzy  $a$  i  $c$  (zgodnie z przedstawionym powyżej schematem). Natomiast istnienie związków pomiędzy  $a$  i  $b$  zostało tylko założone. Obecnie spróbujemy założenia te zweryfikować, ponieważ będzie to ważne dla budowy teorii dotyczącej procesów wewnętrznych respondenta w trakcie udzielania odpowiedzi na pytania kwestionariuszowe.

<sup>4</sup> M. Choynowski, *Elementy teorii testów psychologicznych*, „Przegląd Psychologiczny”, 1959, nr 3; D. Magnuson, *Test Theory*, 1967.

<sup>5</sup> Patrz Lutyński, *Analizy weryfikacyjne...*

W pierwszym założeniu przyjęto, że respondent o małym zasobie słów będzie miał trudności ze zrozumieniem pytań. Nie budzi ono sprzeciwu, warto jednak poprzeć je empirycznym dowodem. Wykorzystując materiały z poprzednich badań, można obliczyć średnie wyniki w teście słownikowym osób, które rozumiały większą część pytań, osób, które nie rozumiały całkowicie co najmniej jednego pytania (tzn. nie rozumiały terminów zawartych w pytaniu i dopiero dodatkowe wyjaśnienia spowodowały dobre rozumienie pytania), osób, które nie rozumiały co najmniej jednego pytania i mimo wyjaśnień ze strony ankietera nie mogły zrozumieć nadal pytania.

Gdyby założenie przyjęte w poprzednim badaniu było słuszne, to średnie wyniki w teście słownikowym w poszczególnych grupach powinny maleć zgodnie z podaną kolejnością (hipoteza I).

Podobnie, można sprawdzić empirycznie inne założenia. Przyjęto poprzednio, że osoby uzyskujące wysokie wyniki w teście „kłamstwa” mają tendencję do pokazywania siebie w lepszym świetle i dlatego dają odpowiedzi społecznie akceptowane. Ta skłonność może zmniejszyć ilość udzielanych przez nich prawidłowych odpowiedzi. Analiza potwierdziła, że wyniki w teście K skorelowane są negatywnie z ilością prawidłowych odpowiedzi. Można jednak postawić pytanie: czy wyniki tego testu rzeczywiście związane są z tendencją założoną przez autorów? Próba znalezienia na nie odpowiedzi nie jest tak prosta jak w poprzednim przypadku, ponieważ zrozumienie pytania przez respondenta można było stosunkowo łatwo sprawdzić, natomiast ustalenie, że u kogoś występuje tendencja do pokazywania się w lepszym świetle, jest znacznie trudniejsze. Można jednak, korzystając z poprzednich materiałów, wyszukać osoby, które nie rozumiały pytań, a mimo to udzielały na nie istotnych odpowiedzi<sup>6</sup>, podczas gdy prawidłowe zachowanie respondenta powinno polegać na przyznaniu się do niezrozumienia pytania. Przyczyny udzielania odpowiedzi istotnej w takiej sytuacji mogły być następujące:

— przyznanie się do niezrozumienia pytania stawiało badanych (w ich ocenie) w gorszym świetle w oczach ankietera;

— pomimo niezrozumienia zasadniczych terminów w pytaniu badani mogli domyślać się, która z odpowiedzi istotnych jest lepsza z różnych względów i tę właśnie odpowiedź podawali;

— nie rozumiejąc pytań odpowiadali bez głębszego zastanowienia się nad ich treścią.

---

<sup>6</sup> Odpowiedź istotna to odpowiedź dostarczająca informacji poszukiwanej przez badacza.

Gdyby słuszne było założenie o związku wysokich wyników testu *K* z tendencją do pokazywania się w lepszym świetle, wówczas średni wynik w tym teście osób, które udzielają odpowiedzi istotnej bez zrozumienia pytania, powinien być wyższy niż średni wynik osób, które nie rozumiejąc pytania dawały odpowiedź nieistotną; np. przyznawały się, że nie rozumieją pytania, podawały, że trudno im odpowiedzieć na dane pytanie (hipoteza II).

Ponieważ u podstaw dawania odpowiedzi istotnej przy braku zrozumienia pytania leżeć może jeszcze trzecia z wymienionych przyczyn, można będzie sprawdzić słuszność założenia, że osoby skłonne do udzielania odpowiedzi na pytania „puste” zgodnie z żądaniem zawartym w tych pytaniach odpowiadać będą bez głębszego zastanawiania się nad treścią pytania. Gdyby tak było, to osoby udzielające odpowiedzi istotnej przy braku rozumienia pytania powinny również częściej „dawać się łapać” na pytania „puste” niż osoby, które dawały odpowiedzi nieistotne w przypadku niezrozumienia pytania (hipoteza III).

Weryfikacja trzech wymienionych hipotez dostarczy informacji o trafności teoretycznej testów zastosowanych do oceny respondentów odpowiadających na pytanie kwestionariuszowe. Jednocześnie może stanowić ona podstawę do oceny tego, czy założone właściwości mierzone przez te testy odgrywają rolę przy udzielaniu prawidłowych odpowiedzi.

Następne zagadnienie, któremu poświęcona zostanie część niniejszego opracowania, związana jest tylko z testem słownikowym. Autorzy przygotowujący test *S* wyszli z założenia, że powinien on zawierać terminy występujące w pytaniach kwestionariuszowych. Ponieważ test ten okazał się dobrym predykatorem udzielania przez respondentów prawidłowych odpowiedzi, należałoby wyciągnąć wniosek, że test słownikowy powinien być każdorazowo dostosowany do poszczególnych badań. Z praktycznego punktu widzenia jest to duża niedogodność, ponieważ przygotowanie nowego testu wymaga oddzielnych, dość pracochłonnych badań. Wygodniejsze dla badaczy byłoby korzystanie z jednego, odpowiednio przygotowanego testu słownikowego, bez względu na to jakiej problematyki dotyczą badania kwestionariuszowe. Należałoby zatem rozstrzygnąć, czy trzeba, czy nie trzeba dostosowywać test słownikowy do konkretnych badań.

Aczkolwiek założenie przyjęte w poprzednim opracowaniu, że test powinien zawierać terminy występujące w kwestionariuszu, wydaje się słuszne, to jednak dwie racje przemawiają za tym, aby testu nie dostosowywać do poszczególnych badań kwestionariuszowych:

— można przypuszczać, że osoba, która ma duży zasób słownictwa (tzn. uzyskuje dobry wynik w teście słownikowym), będzie również

znała więcej terminów zawartych w pytaniach kwestionariuszowych bez względu na ich tematykę;

— ponadto badacz sam musi dbać o to, aby w tworzonych przez niego pytaniach nie znajdowały się terminy zbyt trudne lub znane tylko wąskiej grupie osób.

Problem ten musi być rozstrzygnięty w sposób empiryczny. Ponieważ zastosowany w badaniach test słownikowy zawierał terminy związane i nie związane z pytaniami kwestionariuszowymi, można będzie wyodrębnić dwa podtesty. Jeden podtest będzie zawierał terminy związane, a drugi nie związane z pytaniami. Następnie obliczona zostanie korelacja pomiędzy tymi dwoma podtestami oraz korelacja pomiędzy każdym podtestem osobno a liczbą prawidłowych odpowiedzi udzielanych przez badanych.

Za tym, że testu nie trzeba dostosowywać każdorazowo do badań kwestionariuszowych, powinny przemawiać następujące wyniki:

— wysoka korelacja pomiędzy podtestami, która świadczyłaby o ich równoważności,

— co najmniej podobne współczynniki korelacji pomiędzy podtestami a ilością udzielanych prawidłowych odpowiedzi.

Natomiast niska korelacja pomiędzy podtestami, a przede wszystkim znacznie wyższa korelacja pomiędzy podtestem zawierającym terminy związane z pytaniami kwestionariuszowymi a liczbą prawidłowych odpowiedzi niż korelacja drugiego podtestu z tym kryterium, powinna być dowodem na to, że trzeba dostosowywać test słownikowy do poszczególnych badań kwestionariuszowych.

#### ANALIZA WYNIKÓW

Badano 122 pracowników fizycznych. Tylko w analizach korelacyjnych uwzględniono wyniki testowe wszystkich badanych. Natomiast pozostałe analizy oparte są na wynikach mniejszej liczby osób wybranych z całej grupy ze względu na weryfikowaną hipotezę.

Hipoteza I brzmi następująco: średnie wyniki w teście słownikowym osób, które w większości rozumiały pytania kwestionariuszowe (grupa A); osób, które nie rozumiały co najmniej jednego pytania (grupa B); oraz osób, które nie rozumiały co najmniej jednego pytania i mimo wyjaśnień nie mogły go zrozumieć (grupa C); będą coraz to mniejsze, zgodnie z podaną kolejnością.

Tabela 1 zawiera wyniki potrzebne do weryfikacji tej hipotezy. Układają się one zgodnie z przyjętą hipotezą. Różnica pomiędzy średnim wynikiem grupy A a średnimi wynikami pozostałych grup jest

statystycznie istotna ( $p < 0,01$ ). Natomiast różnica pomiędzy średnimi grup B i C jest statystycznie nieistotna.

Tabela 1. Średnie wyniki w teście słownikowym osób z grup A, B i C

| Grupa | N  | $\bar{x}$ | $\sigma$ |
|-------|----|-----------|----------|
| A     | 25 | 19,7      | 4,9      |
| B     | 40 | 8,7       | 3,2      |
| C     | 12 | 7,5       | 2,6      |

Tabela 2. Średnie wyniki w teście „kłamstwa” grupy D i E

| Grupa | N  | $\bar{x}$ |
|-------|----|-----------|
| D     | 42 | 19,4      |
| E     | 10 | 18,1      |

Tabela 2 przedstawia wyniki umożliwiające weryfikację hipotezy II mówiącej, że osoby, które co najmniej raz udzielają odpowiedzi istotnej bez zrozumienia treści pytania (grupa D), uzyskają przeciętnie wyższe wyniki w teście „kłamstwa” niż osoby, które nie rozumiejąc pytania dawały zawsze odpowiedzi nieistotne (grupa E). Kierunek różnicy średnich jest zgodny z postawioną hipotezą, lecz różnica pomiędzy rozkładami obu grup okazała się nieistotna. Ponieważ rozkłady wyników w obu grupach odbiegały od normalnego oraz ich wariancje nie były podobne, istotność różnicy rozkładów sprawdzono testem serii  $u$ .

Hipoteza III stwierdzała, że osoby, które udzielają odpowiedzi istotnej na pytania kwestionariuszowe bez zrozumienia ich znaczenia (grupa D), częściej udzielają odpowiedzi istotnej na pytania „puste” niż osoby, które dawały odpowiedzi nieistotne przy braku zrozumienia pytania kwestionariuszowego (grupa E). Za osoby udzielające częściej odpowiedzi istotnej na pytania „puste” uznano te osoby, które odpowiedziały w ten sposób najmniej na 4 z 5 pytań (takie samo kryterium podziału zastosowano w poprzednich badaniach, które dostarczały materiału do obecnej analizy).

Tabela 3. Liczba i procent osób z grupy D i E, które udzielały rzadziej i częściej odpowiedzi istotnych na pytania „puste”

| Częstość udzielania odpowiedzi istotnej na pytania „puste” | Grupa |     |    |     |
|--|-------|-----|----|-----|
|  | D     |     | E  |     |
|  | N     | %   | N  | %   |
| 0—3  | 4     | 9   | 2  | 20  |
| 4—5  | 38    | 91  | 8  | 80  |
| Razem  | 42    | 100 | 10 | 100 |

Tabela 3 przedstawia uzyskane wyniki. Wynika z niej, że osoby z grupy D dawały częściej odpowiedzi istotne („dawały się łapać”) na pytania „puste”, co zgodne jest z postawioną hipotezą. Jednak różnice pomiędzy obu grupami są statystycznie nieistotne.



Następna część analizy dotyczy pytania: czy trzeba dostosowywać test słownikowy do poszczególnych badań kwestionariuszowych? Odpowiedzi na to pytanie dostarczyć miały analizy korelacyjne pomiędzy podtestem zawierającym terminy występujące w kwestionariuszu ( $S_k$ ) a podtestem składającym się z reszty zadań ( $S_r$ ) oraz pomiędzy wynikami uzyskanymi przez osoby badane w wymienionych podtestach a liczbą prawidłowych odpowiedzi na pytania kwestionariuszowe (zmienna  $x$ ).

Uzyskano następujące współczynniki korelacji:

$$r_{S_k S_r} = +0,65, \quad (1)$$

$$r_{S_r x} = +0,58, \quad (2)$$

$$r_{S_k x} = +0,54. \quad (3)$$

Współczynniki te wskazują, że korelacja pomiędzy dwoma podtestami jest dość wysoka. Natomiast korelacje pomiędzy wynikami uzyskanymi w poszczególnych podtestach z liczbą prawidłowych odpowiedzi (co stanowi kryterium jakości respondenta) są do siebie bardzo zbliżone. Różnice pomiędzy nimi są statystycznie nieistotne. Świadczy to o tym, że trafność z kryterium obu podtestów jest taka sama. Na obniżenie trzeciego współczynnika korelacji wpłynął zapewne fakt, że podtest ten składał się z małej ilości zadań.

#### DYSKUSJA I WNIOSKI

1. Pomimo braku statystycznej istotności różnicy pomiędzy średnimi wynikami testowymi grup *B* i *C* (tab. 1) można przyjąć, że hipoteza I została potwierdzona. Osoby, które nie rozumiały pytań kwestionariuszowych, uzyskiwały znacznie gorsze wyniki w teście słownikowym. Z tego wypływa wniosek, że test ten bada rozumienie pytań. Jest to dodatkowy dowód wskazujący na przydatność testu słownikowego do wartościowania respondentów pod kątem udzielania przez nich prawidłowych odpowiedzi.

2. Dwie następne hipotezy nie uzyskały statystycznego potwierdzenia, lecz wyniki układają się zgodnie z tymi hipotezami. Brak istotnej różnicy pomiędzy wynikami testowymi grup *D* i *E* (tab. 2 i 3) mógł wynikać z tego, iż grupa *E* była zbyt mało liczna. Po drugie, tendencja do pokazywania się w lepszym świetle, którą mierzy test „kłamstwa” różni się od tej tendencji, która występuje w sytuacji, gdy badani nie rozumiejąc pytań dają na nie odpowiedzi istotne. Badani odpowiadając na pytania w teście „kłamstwa” w zasadzie je rozumieją, bo są one łatwe, czyli gdy odpowiadają na nie w sposób, który pokazuje ich

w lepszym świetle w oczach ankietera, czynią to kontrolując intelektualnie treść pytania. Natomiast w drugim przypadku badani nie rozumieją pytań, a mimo to odpowiadają tak, jak gdyby je rozumieli. Mogło to być spowodowane chęcią uniknięcia wstydu, iż nie potrafią zrozumieć, co się do nich mówi.

Nie wyjaśnia to całkowicie braku potwierdzenia hipotezy II, bo gdyby wyciągnąć konsekwencje z powyższego rozumowania, należałoby oczekiwać wyraźnej zależności pomiędzy udzielaniem przez badanych odpowiedzi istotnych zarówno na pytania „puste”, jak i na niezrozumiałe. Takiej zależności jednak nie stwierdzono (patrz tab. 3).

Najprawdopodobniej niewielkie różnice w wynikach obu testów pomiędzy grupami *E* i *D* są rezultatem tego, że do grup osób, które dawały odpowiedzi istotne przy braku rozumienia pytania weszli ci badani, którzy

— udzielali odpowiedzi istotnych, chcąc świadomie wypaść lepiej w oczach ankietera;

— nie przyznawali się do niezrozumienia pytania, chcąc uniknąć wstydu;

— udzielali odpowiedzi bez głębszego zastanowienia się nad treścią pytań, czyli odpowiadali w sposób „automatyczny”.

Mogło to spowodować niejednorodność grupy pod względem motywacji kierującej procesem odpowiadania na pytania kwestionariuszowe. Można przypuszczać, na podstawie analizy treści pytań testu „kłamstwa”, że mierzy ona raczej pierwszą tendencję i w pewnym stopniu drugą, a test pytań „pustych” mierzy raczej trzecią tendencję. Tak więc niejednorodność grupy mogła przyczynić się do braku istotności różnic ujawnionych w tabelach 2 i 3.

Dwa wnioski wypływają z tego dla przyszłych badań. Należy przebadąć większą grupę osób, a następnie przy pomocy odpowiednich metod statystycznych ustalić poszczególne związki wymienianych powyżej czynników, które spowodowały niejednorodność grupy, z wynikami testów „kłamstwa” i pytań „pustych”. Trafność teoretyczną testu „kłamstwa” można będzie ustalić sprawdzając, czy osoby, które są przez ich kolegów oceniane jako posiadające tendencję do pokazywania się w lepszym świetle, rzeczywiście uzyskują w tym teście wyższe wyniki niż osoby nie posiadające tej tendencji.

3. Analizy korelacyjne przeprowadzone pomiędzy podtestami wydzielonymi z testu słownikowego a miarą kryterium jakości respondenta (liczba udzielonych prawidłowych odpowiedzi) wykazały, że nie trzeba każdorazowo dostosowywać testu do terminów występujących w pytaniach kwestionariuszowych. Oba te podtesty posiadają podobny wskaźnik trafności.