

Anna PRUSAK^{1*}
Piotr STEFANÓW^{2**}
Magdalena GARDIAN³

GRAFICZNA FORMA KWESTIONARIUSZA W BADANIACH AHP/ANP

AHP, czyli Analityczny Proces Hierarchiczny oraz jej rozszerzenie ANP, czyli Analityczny Proces Sieciowy to jedne z najbardziej znanych w świecie wielokryterialnych metod wspomaganie decyzji. Gromadzenie danych w metodach AHP/ANP polega na dokonaniu przez eksperta porównań parami między sobą wszystkich elementów modelu hierarchicznego bądź sieciowego. Celem artykułu jest weryfikacja tezy, że wysoka wartość współczynnika CR (ang. *Consistency Ratio*) w badaniach AHP/ANP, wskazującego na brak logiczności (zgodności) wyników, jest związana z brakiem symetrii osądów wynikającego z niewłaściwie zaprojektowanego kwestionariusza. Zaprojektowano eksperyment, którego celem było sprawdzenie, czy frakcja odpowiedzi powyżej i poniżej wartości środkowej jest probabilistycznie jednakowa, a jeśli nie, to czy miało to wpływ na wysoką wartość współczynnika CR. Postawiona hipoteza stanowi, że frakcja, obliczana jako iloraz odpowiedzi powyżej wartości środkowej do odpowiedzi leżących powyżej i poniżej wartości środkowej, powinna wynosić w przybliżeniu $\frac{1}{2}$. Do weryfikacji hipotezy o symetrycznym rozkładzie odpowiedzi wykorzystano test serii. W kilku badaniach, w których przeprowadzono w sumie 8434 porównań, wzięło udział ponad 400 respondentów poproszonych o ocenę różnych obiektów. Kolejność porównywanych obiektów oraz liczba kryteriów oceny była różna i ułożona losowo. Do eksperymentu wykorzystano 9-stopniową, dwubiegunową skalę werbalną w formie pionowej tabelarycznej. W prawie każdym przypadku uzyskano wynik wskazujący, że symetria odpowiedzi została zachowana. Stąd też można uznać, że ankieta została zaprojektowana w sposób właściwy, a źródła wysokiego poziomu niezgodności należy szukać gdzie indziej. Przedstawiony sposób weryfikacji kwestionariusza ankiety nie dotyczy tylko metody AHP/ANP, a zaniechanie sprawdzenia „jakości” kwestionariusza może prowadzić do sytuacji, w której uzyskane wyniki będą obarczone błędami.

Słowa kluczowe: analityczny proces hierarchiczny i sieciowy, AHP, ANP, zgodność, kwestionariusz, ankieta

1. WSTĘP

Metody AHP, czyli Analityczny Proces Hierarchiczny (ang. *Analytic Hierarchy Process*), oraz jej rozszerzenie ANP, czyli Analityczny Proces Sieciowy (ang. *Analytic Network Process*) to jedne z najbardziej znanych w świecie wielokryterialnych metod wspomaganie decyzji (ang. *MCDM, Multiple Criteria Decision Analysis*). Zostały stworzone w latach 70 XX w. przez T.L. Saaty'ego, amerykańskiego profesora matematyki. Przykładem innych metod i technik wspomagających podejmowanie decyzji są na przy-

¹ Dr Anna Prusak, adres: Uniwersytet Ekonomiczny w Krakowie, Katedra Zarządzania Jakością, ul. Rakowicka 27, 31-510 Kraków, e-mail: anna.prusak@uek.krakow.pl, telefon: 12-293-75-16 (Autor korespondencyjny).

² Piotr Stefanów, Krakowska Akademia im. A.F. Modrzewskiego, e-mail: piotr@stefanow.net.

³ Magdalena Gardian, ul. Chocimska 39/9, Kraków, e-mail: magda.gardian@gmail.com.

kład ELECTRE (fr. *ELimination Et Choix Traduisant la REalite*), PROMETHEE & GAIA (ang. *Preference Ranking Organization Method for Enrichment of Evaluations & Geometrical Analysis for Interactive Aid*), MACBETH (ang. *Measuring Attractiveness by a Categorical Based Evaluation Technique*) czy VDA (ang. *Verbal Decision Analysis*). Jednakże żadna z wyżej wymienionych metod nie zdobyła takiej popularności i mocnej pozycji w światowej nauce, jak metody AHP/ANP. Świadczy o tym m.in. duża i stale rosnąca liczba publikacji nt. AHP/ANP w międzynarodowych bazach wysoko punktowanych czasopism recenzowanych, jak np. *ScienceDirect*, *Wiley*, *Scopus* [Prusak i Stefanów 2011]. Oprócz szerokiego zastosowania w badaniach naukowych, metody AHP/ANP są od wielu lat wykorzystywane przez menedżerów dużych organizacji (zarówno rządowych, jak i prywatnych) w rozwiązywaniu bieżących problemów decyzyjnych. Przykładem takich organizacji są Departament Obrony USA, IBM, British Airways, Xerox, Ford i wiele innych [Saaty 2008]. Metody AHP/ANP wyróżniają się przede wszystkim wszechstronnością, elastycznością oraz uniwersalnością zastosowania. Nie ma chyba dziedziny, w której nie mogłyby być stosowane, począwszy od ekonomii i zarządzania a skończywszy na naukach medycznych i technicznych. Ich podstawową zaletą jest to, że w stosunkowo prosty sposób pozwalają na dekompozycję złożonych problemów decyzyjnych i ich strukturyzację za pomocą modelu hierarchicznego (AHP) lub sieciowego (ANP), dopuszczając jednoczesną analizę czynników o charakterze zarówno jakościowym, jak i ilościowym. Modele te podlegają następnie systematycznej analizie poprzez dokonanie porównań parami poszczególnych ich elementów, oraz obliczeniu odpowiednich współczynników wagowych. W ten sposób warianty decyzyjne porządkowane są od najbardziej do najmniej preferowanego, tym samym wskazując kierunek ostatecznej decyzji. Metod tych używa się do analizy różnych form problemów decyzyjnych, m.in. alokacja zasobów, wyznaczanie strategii, prognozowanie, rozwiązywanie konfliktów. Mogą być one łączone z innymi wielokryterialnymi metodami i technikami wspomagania decyzji, jak np. wyżej wymienione ELECTRE, PROMETHEE, MACBETH i VDA. W literaturze można także często spotkać połączenia typu SWOT-AHP np. [Lee i Walsh 2011], AHP/DEA np. [Sinuany-Stern, Mehrez i Hadad 2000], Delphi-AHP, np. [Joshi, Benwet i Shankar 2011]. Dzięki uzasadnieniu wyników za pomocą odpowiednich obliczeń, metody te eliminują ryzyko manipulacji podczas dokonywania decyzji. Same obliczenia mogą zostać zrobione w miarę szybko, za pomocą dedykowanego oprogramowania komputerowego (np. *Super Decisions*, *Decisions Lens*, *Expert Choice*). Dodajmy, że omawiane metody spotkały się także z krytyką, która dotyczy przede wszystkim tzw. niezgodności (nielogiczności) rezultatów. Wynika ona z wysokiej wrażliwości AHP/ANP na nawet najdrobniejsze „nieścisłości” i błędy powstałe na etapie gromadzenia danych.

Gromadzenie danych w metodach AHP/ANP polega na dokonaniu porównań parami między sobą wszystkich elementów modelu hierarchicznego bądź sieciowego. Do pomiaru poziomu niezgodności wyników służy specjalny współczynnik zgodności CR (ang. *Consistency Ratio*). Aby wyniki można było uznać za zgodne, wartość tego współczynnika nie może przekraczać 0,10 (10%), co jest szczególnie trudne do osiągnięcia w przypadku dużej liczby elementów. Co więcej, w niektórych przypadkach przyjmuje się jeszcze niższy dopuszczalny poziom CR, którym jest $CR=0,05$ dla $n=3$ elementów [Saaty 1994]. Należy przy tym podkreślić, że sama możliwość pomiaru niezgodności za pomocą odpowiedniej procedury obliczeniowej powinna być postrzegana jako zaleta metody AHP, gdyż pozwala na sprawdzenie, czy respondent odpowiadał na zadane pytania w sposób przemyślany, a nie przypadkowy. Na poziom tej przypadkowości odpowiedzi

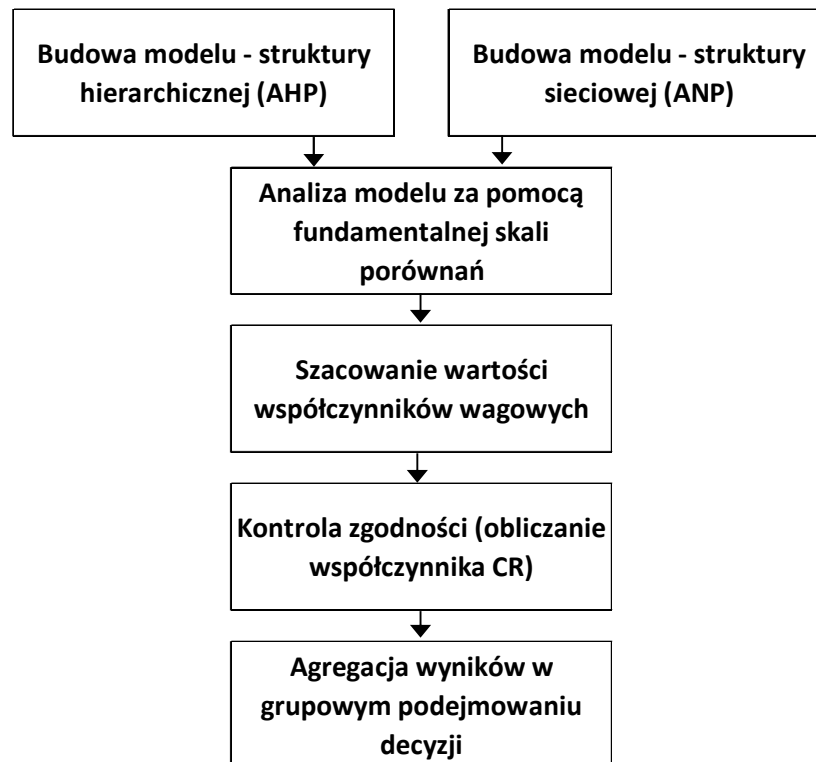
wpływa wiele czynników, a jednym z nich jest graficzna forma kwestionariusza. Dotychczasowe badania np. [Webber, Apostolou i Hassel 1996, Weathers, Sharmab i Niedricha 2005] oraz obserwacje dokonane podczas prowadzenia badań⁴ przez autorów pokazują, że graficzna forma ankiety wpływa na sposób udzielanych odpowiedzi. Na przykład, w przypadku 9-stopniowej, dwubiegunowej skali używanej w metodach AHP/ANP respondenci mają tendencję do „automatycznego” udzielania odpowiedzi tylko po jednej stronie skali w niektórych formach graficznych ankiety. Ponieważ wiązało się to z wysoką wartością współczynnika CR, wysunięto tezę, że zachowanie symetrii odpowiedzi jest warunkiem koniecznym (ale niewystarczającym) zachowania logiczności osądów. Celem prezentowanych badań jest weryfikacja tej tezy, polegająca na zbadaniu symetrii odpowiedzi w ankietach AHP przy wykorzystaniu statystycznego testu symetrii. Do badań użyto jedną z najczęściej stosowanych form prezentacji graficznej kwestionariusza: dwubiegunową skalę werbalną z punktem środkowym w układzie pionowym. Przeprowadzony eksperyment pozwolił wykluczyć hipotezę, że wysoki odsetek niezgodnych odpowiedzi jest wynikiem braku symetrii osądów wynikającego ze źle zaprojektowanego kwestionariusza AHP/ANP. Dlatego w tym przypadku przyczyn wysokiego CR należy szukać gdzie indziej.

W kolejnej części artykułu zostaną objaśnione poszczególne etapy metody AHP, na podstawie której będzie przeprowadzony eksperyment. Warto jednak podkreślić, że wyniki odnosić się będą w takim samym stopniu do metody ANP, w której używa się tej samej skali i kwestionariusza. Następnie zostanie przedstawiona charakterystyka użytego do badań narzędzia, czyli kwestionariusza AHP. Zostaną zaprezentowane jego najczęściej stosowane formy graficzne. W kolejnej sekcji będzie omówiony eksperyment badawczy oraz jego wyniki.

2. ETAPY METODY AHP

Etapy analizy AHP/ANP oraz ich fundamenty matematyczne można znaleźć praktycznie w każdym podręczniku i artykule T. Saaty’ego oraz w większości publikacji dotyczących tych metod. W sposób ogólny zostały one przedstawione na poniższym schemacie (Rysunek 1).

⁴ Badania były prowadzone w ramach grantu MNiSW NN111345138: *Metodologia wielokryterialnej analizy porównawczej obiektów*.



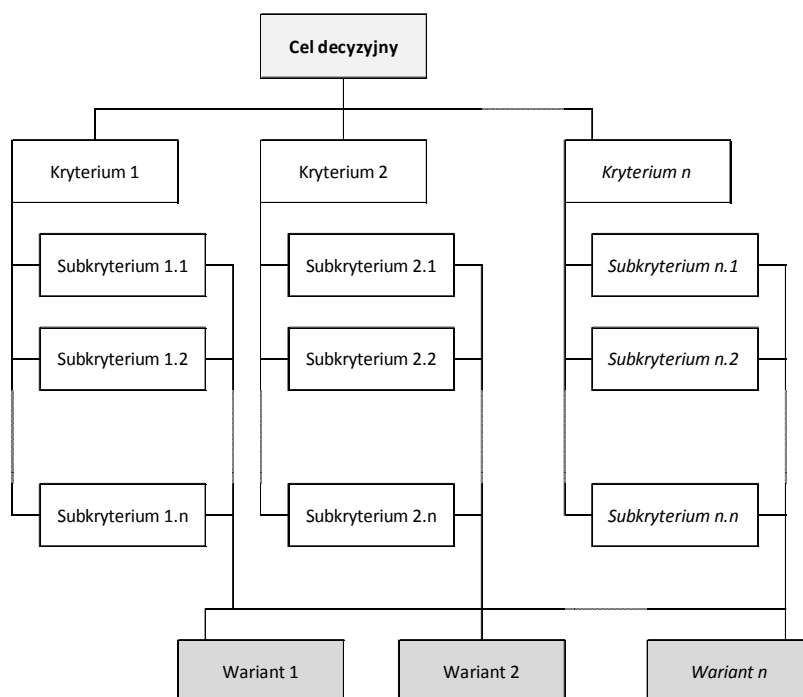
Rysunek 1. Etapy metody AHP

Figure 1. Stages of the AHP method

Źródło: opracowanie własne

Etap 1: Budowa modelu decyzyjnego

Konstrukcja modelu decyzyjnego jest pierwszym etapem analizy AHP lub ANP. Jak wcześniej wspomniano, obecny artykuł koncentruje się na metodzie AHP, w przypadku której model decyzyjny przyjmuje postać struktury hierarchicznej. Przykład najprostszej, najczęściej stosowanej struktury hierarchicznej został przedstawiony na Rysunku 2. Jest to czteropoziomowy model złożony z celu decyzyjnego („co chcemy osiągnąć?”) umieszczonego na samym szczycie hierarchii, kryteriów decyzyjnych („co ma wpływ na realizację celu decyzyjnego?”), tzw. sub-kryteriów („co ma wpływ na spełnienie poszczególnych kryteriów?”) oraz wariantów decyzyjnych („jaki mamy wybór?”), zwanych również „alternatywami” decyzyjnymi i mającymi miejsce na samym dole modelu. Istotą dalszej analizy tego modelu będzie wskazanie tego wariantu, który w najwyższym stopniu spełnia najważniejsze sub-kryteria oraz kryteria, a tym samym przyczynia się w najwyższym stopniu do realizacji celu decyzyjnego.



Rysunek 2. Czteropoziomowy model hierarchiczny

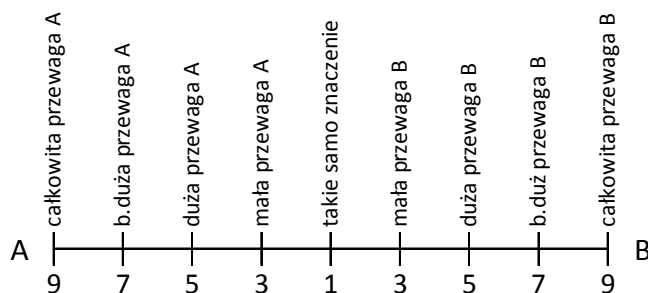
Figure 2. Four-level hierarchical model

Źródło: opracowanie własne

Analiza tak zbudowanego modelu hierarchicznego ma miejsce w układzie poszczególnych grup elementów, zawsze z poziomu niższego względem poziomu znajdującego się bezpośrednio wyżej. Na przykład, kryteria analizowane są względem celu, subkryteria względem odpowiadających im kryteriów, natomiast warianty decyzyjne w stosunku do każdego subkryterium. Modele sieciowe stanowią rozszerzenie modeli hierarchicznych, gdyż pozwalają uwzględnić inne typy relacji zachodzących pomiędzy poszczególnymi elementami, w tym pozwalają na analizę elementów pomiędzy grupami i we wszystkich kierunkach. Kolejne etapy przebiegają w ten sam sposób zarówno dla metody AHP jak i ANP, w oparciu o te same zasady, dlatego do realizacji celu obecnej pracy użyto jako przykładu modelu hierarchicznego. Większa złożoność modeli sieciowych i konieczność wykonania dodatkowych obliczeń nie wpływa na wnioski wynikające z przedstawionych badań empirycznych.

Etap 2: Analiza modelu za pomocą fundamentalnej skali porównań

Do analizy modelu decyzyjnego używa się specjalnej skali, zwanej fundamentalną, bądź 9-stopniową skalą porównań. Skala ta jest praktycznie jedyną skalą zalecaną przez twórcę metod AHP/ANP do analizy modeli, oparte jest na niej także oprogramowanie komputerowe służące do analizy AHP/ANP. Jest to skala liniowa, dwubiegunowa, służąca do porównań parami różnych obiektów, czyli do ustalania stopnia przewagi jednego elementu nad drugim. Mając dwa elementy A i B posiadające jakąś wspólną właściwość, respondent wskazuje, który z nich ma względem tej właściwości przewagę, i jak duża jest ta przewaga (Rysunek 3).



Rysunek 3. 9-stopniowa fundamentalna skala porównań

Figure 3. Nine-point fundamental comparison scale

Źródło: opracowanie własne

Wielkość przewagi jednego elementu nad drugim można określić zarówno werbalnie, jak i liczbowo, przy czym zaleca się, aby respondentom dokonującym analizy przedstawiać raczej formę słowną niż liczbową, według następujących stopni przewagi:

- A i B mają takie samo znaczenie („1”),
- A ma małą przewagę nad B lub B ma małą przewagę nad A („3”),
- A ma dużą przewagę nad B lub B ma dużą przewagę nad A („5”),
- A ma bardzo dużą przewagę nad B lub B ma bardzo dużą przewagę nad A („7”),
- A ma całkowitą przewagę nad B lub B ma całkowitą przewagę nad A („9”).

Stopnie parzyste znajdujące się „pomiędzy” głównymi kategoriami (oznaczone jako 2, 4, 6, 8) określane są jako „wartości pośrednie” i respondent powinien je stosować tylko wówczas, gdy nie jest w stanie dokonać rzetelnej oceny posługując się jedynie „głównymi” stopniami przewagi. Aby struktura była zanalizowana w sposób poprawny, każdy element musi być porównany z innym, znajdującym się w tej samej grupie modelu. Liczbę tak utworzonych par elementów można policzyć korzystając ze wzoru:

$$N = \frac{n(n-1)}{2} \quad (1)$$

gdzie:

n jest liczbą porównywanych elementów.

Jak łatwo można zauważyć, duża liczba elementów skutkuje dużą liczbę porównań, co w konsekwencji niesie za sobą ryzyko pomyłki wynikającej ze stanu koncentracji odpowiadającego. Na przykład, $n=8$ elementów będzie wymagać dokonania porównań między sobą w aż 28 kombinacjach, a dla $n=10$ już 45. Zaleca się, aby liczba elementów w obrębie grup modelu wynosiła co najwyżej 9. Zostało one sformułowane na podstawie tzw. magicznej liczby 7 ± 2 Millera. Zgodnie z tą regułą, człowiek nie jest w stanie dokonać w tym samym czasie w sposób logiczny analizy więcej niż 9 elementów [Miller 1956]. Stopień logiczności dokonanych porównań jest mierzony za pomocą specjalnego współczynnika zgodności CR (ang. *Consistency Ratio*), po wprowadzeniu wyników porównań do macierzy i obliczeniu współczynników wagowych. Ponieważ współczynnik CR jest bardzo wrażliwy na logiczność dokonywanych osądów, obecny artykuł koncentruje się na zbadaniu sposobu udzielania odpowiedzi przez respondentów po to, aby możliwe było zaproponowanie takiego sposobu prowadzenia badań, który pozwoli uniknąć błędów prowadzących do dużych niezgodności.

Etap 3: Szacowanie wartości współczynników wagowych

Wyniki porównań wprowadzane są do kwadratowej macierzy A o wymiarach $(n \times n)$, gdzie n jest liczbą porównywalnych elementów. Macierz ta stanowi główne narzędzie analizy metody AHP np. [Saaty 2006]:

$$A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \\ 1/a_{12} & 1 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \dots & 1 \end{bmatrix} \quad (2)$$

Elementy leżące na przekątnej tej macierzy są zawsze jedynkami, gdyż:

$$a_{ij} = 1, \quad \text{dla} \quad i = j \quad (3)$$

Jest to macierz odwrotnie symetryczna, ponieważ zgodnie z regułą odwrotności:

$$\frac{1}{a_{ij}} = a_{ji} \quad (4)$$

Na przykład, jeśli element A ma dużą przewagę nad B czyli $A=5B$, to $B=1/5A$. Macierzy buduje się tyle, ile w modelu znajduje się grup elementów. Po wykonaniu określo-

nych operacji matematycznych na macierzy otrzymuje się tzw. wektory priorytetów W (w_1, w_2, w_3), które określane są także jako współczynniki wagowe (wagi, priorytety) i które wskazują relatywną ważność porównywanych elementów. Współczynniki wagowe można obliczyć za pomocą kilku sposobów. Najprościej jest skorzystać z oprogramowania komputerowego, a w przypadku braku dostępu do niego można przeprowadzić „ręczną” kalkulację za pomocą jednej z procedur tj. mnożenia macierzy, wykorzystując średnią arytmetyczną lub średnią geometryczną. Najczęściej omawiane w literaturze są metody średniej arytmetycznej lub geometrycznej, które jednak generują jedynie przybliżone wyniki. Dokładne omówienie procedur obliczania współczynników wagowych można znaleźć w bogatej literaturze przedmiotu np. [Saaty 994], [Ishizaka i Labib 2009].

Etap 4. Kontrola zgodności porównań (obliczanie współczynnika CR)

Dla każdej macierzy obliczany jest następnie współczynnik zgodności CR, który wskazuje, jak logiczne (zgodne) są oceny dokonane przez respondentów, a zatem jak wiarygodne będą otrzymane na ich podstawie wyniki. Przyjmuje się, że aby uznać wyniki za zgodne, wartość współczynnika CR nie może przekroczyć 0,10 (10%). Jeśli CR przekroczy ten poziom, porównania należy uznać za niezgodne, a analizę należy powtórzyć np. [Saaty i Vargas 1982]. Współczynnik ten można obliczyć zarówno przy użyciu programu komputerowego, jak i „ręcznie”. Podstawą jest wyznaczenie tzw. największej wartości własnej (λ_{\max}) macierzy A . W uproszczony sposób obliczana jest ona jako suma iloczynów sumy wartości porównań w każdej kolumnie macierzy A oraz odpowiedniego dla danego elementu współczynnika wagowego:

$$\lambda_{\max} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \cdot w_i \quad (5)$$

Następnie oblicza się tzw. indeks niezgodności (*Inconsistency Index – IC*), który jest ilorazem różnicy największej wartości własnej macierzy i liczby porównywanych elementów (n) przez różnicę ($n - 1$):

$$IC = (\lambda_{\max} - n) / (n - 1) \quad (6)$$

Współczynnik CR obliczany jest poprzez podzielenie IC przez stabilizowaną wartość RI:

$$CR = IC / RI \quad (7)$$

Stabilizowane wartości RI (ang. *Random Index*) zostały wygenerowane przez Saaty’ego na podstawie symulacji dla 500 000 macierzy, i ich wartości zostały przedstawione w Tabeli 1.

Tabela 1: Stablicowane wartości RI

Table 1: Table of RI values

<i>n</i>	2	3	4	5	6	7	8	9	10	11	12
<i>R</i>	0,00	0,52	0,89	1,11	1,25	1,35	1,40	1,45	1,49	1,52	1,54

Źródło: Opracowanie własne na podst. [Alonso i Lamata 2005]

Ponieważ w przypadku wysokiej wartości CR powtarzanie analizy jest bardzo czasochłonne, zaproponowano stosowanie pewnych algorytmów, które mają na celu zredukowanie wartości tego współczynnika, np. „algorytm genetyczny” [Costa, Wanderley i Cosenza 2006]. Takie postępowanie zawsze narusza jednak informacje wejściowe, dlatego najlepiej uzyskać zgodne wyniki już na etapie gromadzenia danych. Warto dodać, że oprócz kontroli zgodności należy również przeprowadzić analizę wrażliwości, która pozwala na zbadanie wpływu zmiany wartości macierzy na wyniki końcowe współczynników wagowych.

Etap 5. Agregacja wyników w grupowym podejmowaniu decyzji

Wielokryterialne metody podejmowania decyzji używane są przede wszystkim do rozwiązywania złożonych i istotnych problemów decyzyjnych. Takie decyzje nigdy nie zapadają jednoosobowo, lecz wymagają udziału wieloosobowych zespołów, złożonych z ekspertów z różnych dziedzin. Wykorzystując metody AHP/ANP stosuje się dwa sposoby matematycznej agregacji ocen w grupowym podejmowaniu decyzji. Są to Agregacja Indywidualnych Osądów (AIJ) realizowana wspólnie przez grupę na etapie porównań parami, oraz Agregacja Indywidualnych Priorytetów (AIP), która polega na syntezie ostatecznych wartości priorytetów pozyskanych od niezależnych respondentów. Więcej na temat decyzji grupowych w metodach AHP/ANP można znaleźć np. w [Saaty i Peniwati 2008]. To zagadnienie nie jest jednak przedmiotem rozważań obecnego artykułu.

3. CHARAKTERYSTYKA UŻYTEGO NARZĘDZIA BADAWCZEGO – KWESTIONARIUSZ AHP

Właściwa forma graficzna oraz wygląd zewnętrzny kwestionariusza ankiety są równie istotne jak (poprawnie metodologicznie) przygotowanie poszczególnych jego elementów [Fanning 2005]. Już samo zapewnienie czytelności i przejrzystości wizualnej formularza powoduje obniżenie błędów popełnianych przez respondentów podczas kodowania odpowiedzi (w tym omijanie pytań), a także wzrost poziomu ich wewnętrznej motywacji do rzetelnego udziału w badaniu. W tym celu zaleca się m.in. stosowanie ujednoliconego graficznie schematu konstrukcji pytań i rozmieszczenia odpowiedzi [Dillman 2000]. Reddline i Dillman [2002] proponują usprawnienie procesu „używania” formularza ankiety poprzez wprowadzenie 4-elementowej ścieżki nawigacyjnej (*navigational path*) – komunikację badacza z respondentem poprzez słowa, cyfry, symbole i grafikę. Zastosowanie odpowiedniej aranżacji graficznej treści może, gdy jest to korzystne, wizualnie wydłużyć lub skrócić dany formularz. Nie należy przy tym jednak pomijać aspektu przejrzystości i czytelności (unikać zbytniego zadrukowania arkuszy). W przypadku stosowania skal graficznych, należy nadać im czytelną formę i zachować jednakowe preferencje wzglę-

dem obu ich biegunów (identyczny wygląd, kolorystyka) [Bradburn, Sudman & Wansink 2004]. Projektując kwestionariusz ankiety uwzględnić należy także liczbę pytań. Poniżej, na Rysunkach 4-9 zaprezentowano przykłady najczęściej stosowanych graficznych form kwestionariusza AHP na przykładzie dwóch czynników wpływających na wybór samochodu: Koszt (K) i Bezpieczeństwo (B). Bardzo często podstawowym kryterium wyboru graficznej formy skali jest ilość miejsca, jaką zajmuje ona w kwestionariuszu. Można zauważyć, że skala werbalna pionowa (Rysunek 4) zajmuje więcej miejsca niż werbalna pozioma (Rysunek 5a i 5b). Kwestionariusz nie powinien być zbyt obszerny, gdyż zbyt duża objętość ankiety powoduje już na początku negatywne nastawienie respondentów do prowadzonych badań. Przykładowe, często wykorzystywane w praktyce, kwestionariusze ankiet zostały zaprezentowane na rysunkach 4-9.

Koszt (K) Który z dwóch czynników ma dla Pana/i większe znaczenie przy zakupie samochodu, i w jakim stopniu? Bezpieczeństwo (B)		(K) jest całkowicie ważniejszy niż (B)
		(K) jest o wiele bardziej ważniejszy niż (B)
		(K) jest dużo ważniejszy niż (B)
		(K) jest niewiele ważniejszy niż (B)
		Obydwa czynniki są tak samo ważne
		(B) jest niewiele ważniejszy niż (K)
		(B) jest dużo ważniejszy niż (K)
		(B) jest o wiele bardziej ważniejszy niż (K)
		(B) jest całkowicie ważniejszy niż (K)

Rysunek 4. Skala werbalna pionowa w formie tabelarycznej.

Figure 4. Verbal scale in vertical tabular form

Źródło: opracowanie własne

kryterium	stopień przewagi										kryterium
	całkowita	bardzo duża	duża	mała (niewielka)	jednakowe znaczenie	mała (niewielka)	duża	bardzo duża	całkowita		
Bezpieczeństwo											Koszt

Rysunek 5a. Skala werbalna pozioma w formie tabelarycznej - wersja (a)

Figure 5a. Verbal scale in horizontal tabular form - version (a)

Źródło: opracowanie własne

kryterium	stopień przewagi								kryterium
	całkowita	bardzo duża	duża	mała (niewielka)	jednakowe znaczenie	mała (niewielka)	duża	bardzo duża	całkowita
Bezpieczeństwo									Koszt

Rysunek 5b. Skala werbalna pozioma w formie tabelarycznej - wersja (b)

Figure 5b. Verbal scale in horizontal tabular form - version (b)

Źródło: opracowanie własne

_____ **Bezpieczeństwo** : **Koszt** _____

Rysunek 6. Skala numeryczna

Figure 6. Numerical scale

Źródło: opracowanie własne

Który z czynników jest ważniejszy przy zakupie samochodu? **W jakim stopniu?**

Bezpieczeństwo _____ **czy** **Koszt** _____

tak samo
ważne średnio dużo bardzo dużo całkowicie

[] [] [] [] [] [] [] [] []

Rysunek 7. Skala werbalna pozioma („dwustopniowa”)

Figure 7. Verbal horizontal scale (“two stages”)

Źródło: opracowanie własne

Czy **Bezpieczeństwo** jest ważniejsze niż **Koszt**?

Jeśli **TAK**, to w jakim stopniu?

całkowicie bardzo dużo dużo średnio tak samo
ważne średnio dużo bardzo dużo całkowicie

[] [] [] [] [] [] [] [] [] [] [] [] [] [] []

Czy **Bezpieczeństwo** jest ważniejsze niż **Koszt**?

Jeśli **TAK**, to w jakim stopniu?

całkowicie bardzo dużo dużo średnio tak samo
ważne średnio dużo bardzo dużo całkowicie

[] [] [] [] [] [] [] [] [] [] [] [] [] [] []

Rysunek 8. Skala werbalna pozioma („jednostopniowa”)

Figure 8. Verbal horizontal scale (“one stage”)

Źródło: opracowanie własne

Który z kryteriów jest ważniejszy przy zakupie samochodu?															
Bezpieczeństwo nad Koszt								czy	Koszt nad Bezpieczeństwo						
								tak samo ważne							
całkowicie	bardzo dużo		dużo		średnio		średnio		dużo		bardzo dużo		całkowicie		
9	8	7	6	5	4	3	2		2	3	4	5	6	7	8

Rysunek 9. Skala mieszana (werbalno-liczbowa)

Figure 9. Mixed scale (verbal and numerical)

Źródło: opracowanie własne

W literaturze dotyczącej metody AHP/ANP brak jest ogólnych zaleceń dotyczących rodzaju i formy użytej skali, poza tym, aby używać raczej skali werbalnej niż liczbowej. Dodatkowo, aby ułatwić respondentowi dokonywanie osądów, należy na początku kwestionariusza zamieścić niezbędne wyjaśnienia wraz z przykładowym, wypełnionym fragmentem ankiety. Zaleca się również, aby badanie AHP było przeprowadzane w asyście moderatora (badacza), który przez cały czas będzie mieć bezpośredni kontakt z respondentem. Wady badań prowadzonych za pomocą samodzielnie wypełnianego kwestionariusza obejmują m.in. brak możliwości interpretacji indywidualnych cech respondenta jak np. gesty i mimika, błędną interpretację lub nadinterpretację pytań i odpowiedzi, brak lub błędnie udzielone (zakodowane) odpowiedzi [Sztumski 1995], małą elastyczność, a także brak możliwości pogłębienia odpowiedzi. Respondent nie może zostać dopytany o rzeczy, które wyszły na jaw w trakcie badania. Jednakże ze względu na fakt, iż w przypadku opisanego eksperymentu istotny był jedynie techniczny sposób wypełniania kwestionariusza przez respondentów i związany z tym odsetek odpowiedzi powyżej i poniżej wartości środkowej, rola badacza ograniczała się do opracowania kwestionariusza, jego dostarczenia i zebrania, a także udzielenia niezbędnych wyjaśnień dotyczących wypełniania ankiety.

W badaniu zastosowano skalę werbalną pionową w formie tabelarycznej (Rysunek 4). Poszczególne pytania nie powtarzały się, a ich liczba była zgodna z założeniami metody AHP. Omawiane zagadnienia oraz poziom ich szczegółowości dobrano adekwatnie do adresatów kwestionariusza, tak aby posiadali oni wystarczające kompetencje do pełnienia roli ekspertów w danym obszarze. Zadbano także, aby pojęcia i sformułowania użyte w kwestionariuszu nie posiadały elementów sugerujących odpowiedź i były wolne od obciążeń (sprawy drażliwe). Ekspert dokonując oceny, zakreślał odpowiednie pole wykorzystując fundamentalną 9-stopniową skalę Saaty'ego. Z uwagi na zapewnienie adekwatności pytań na etapie projektowania eksperymentu nie zastosowano pytań warunkowych (filtrujących). Dodatkowo, w celu uniknięcia zjawiska preferowania przez badanych danego zestawu odpowiedzi występującego przy tabelarycznych (macierzowych) zestawach pytań/ odpowiedzi, oddzielnie przedstawiono poszczególne stwierdzenia tworząc dla nich osobne, choć o identycznym formacie, zestawy odpowiedzi. Aby zminimalizować ryzyko wpływu kolejności pytań na jakość odpowiedzi, ich uszeregowania dokonano w sposób

losowy. Losowe ułożenie pytań było jednak niezwykle istotne przede wszystkim z uwagi na cel prowadzonego eksperymentu. Wygenerowane zestawy (ankiety) różniły się ułożeniem pytań. Zabieg ten był niezbędny dla przeprowadzenia testu symetrii udzielonych odpowiedzi. Każdy respondent po otrzymaniu kwestionariusza, oprócz krótkiego wprowadzenia ze strony badacza, został poproszony do zapoznania się z celem badania i sposobem postępowania. Dodatkowo, załączono dwa przykłady, dla których wyjaśniono interpretację wybranej odpowiedzi oraz przedstawiono jej ilustrację graficzną. Z uwagi na bezpośredni dostęp do ekspertów oraz najwyższy wskaźnik zwrotności kwestionariuszy w eksperymencie [Babbie 2004] przeprowadzono badanie o charakterze audytoryjnym z udziałem ponad 400 respondentów – studentów.

4. OMÓWIENIE I WYNIKI EKSPERYMENTU

Celem eksperymentu było sprawdzenie, czy frakcja odpowiedzi powyżej i poniżej wartości środkowej jest probabilistycznie jednakowa, a jeśli nie, to czy miało to wpływ na wysoką wartość współczynnika CR wskazującego na dużą niezgodność odpowiedzi. Jest to oparte na założeniu, że zachowanie symetrii odpowiedzi jest warunkiem koniecznym, choć niewystarczającym, dla zachowania logiczności osądów. Postawiona hipoteza stanowi, że frakcja, obliczana jako iloraz odpowiedzi powyżej wartości środkowej do odpowiedzi leżących powyżej i poniżej wartości środkowej, powinna wynosić w przybliżeniu $\frac{1}{2}$. Podczas weryfikacji hipotezy o symetrycznym rozkładzie odpowiedzi wykorzystano test serii [Goryl i in. 2007]. Do eksperymentu wykorzystano skalę werbalną pionową w formie tabelarycznej (Rysunek 4).

4.1. Badanie poziomu niezgodności

Badania te miały na celu zbadanie frakcji niezgodnych ocen, mierzonych za pomocą współczynnika CR, w zależności od liczby porównywanych elementów. Każdy z ankietowanych został poproszony o ocenę różnych obiektów. Ich kolejność oraz liczba kryteriów oceny była różna i ułożona losowo. Rolę obiektów oceny pełniły powszechnie dostępne miejsca i przedmioty: mieszkanie, telefon komórkowy, komunikacja miejska, drukarka.

Tabela 2. Obiekty oraz kryteria ich oceny
Table 2. Objects and assessment criteria

Obiekt	Kryteria oceny
Telefon komórkowy [T]	Wielkość wyświetlacza Radio Aparat fotograficzny Wymiary (długość, szerokość) Ciężar telefonu Czas pracy (bateria) Sygnał wibracyjny Wielkość klawiatury

Mieszkanie [M]	Wielkość łazienki Odpowiednie piętro Właściwa technologia (np. cegła) Piwnica Wielkość mieszkania Czas dojazdu do centrum Wielkość kuchni Lokalizacja
Komunikacja miejska [K]	Punktualność Wolne miejsca siedzące Czas oczekiwania na pojazd Cena biletu Brak tłoku Czystość Bezpieczeństwo Szybkość poruszania się
Drukarka [D]	Rozdzielczość Czas wydruku strony Możliwość druku w kolorze Możliwość druku obustronnego Głośność pracy Wymiary Koszty eksploatacji Czas bezawaryjnej pracy
Laptop [L]	Pojemność dysku Szybkość pracy (procesor) Wielkość pamięci operacyjnej Kamera Odpowiednia wielkość ekranu Odpowiedni kolor obudowy Liczba miesięcy gwarancji Ciężar

Źródło: Opracowanie własne

i laptop. Zostały one dobrane w ten sposób, aby ekspertem mogła być dowolna osoba, w tym – ze względu na dostępność respondentów – studenci. Obiekty oraz ich kryteria oceny zostały przedstawione w Tabeli 2.

W każdym przypadku badani byli proszeni o wskazanie, która charakterystyka (kryterium oceny) danego obiektu jest ważniejsza i w jakim stopniu. Ze względu na cel badań oraz możliwość popełnienia podczas badania błędu systematycznego polegającego na tym, że początkowe kryterium „zdeteminuje” kolejne porównania, zdecydowano podczas każdego badania – w sposób losowy – ustalić inną kolejność kryteriów. Eksperyment przeprowadzono w trzech seriach. Najpierw badani oceniali sklep spożywczy [S] ze względu na 3 kryteria. Ponieważ kolejność kryteriów była ustalana w sposób losowy, w efekcie otrzymano 6 różnych wersji ankiety (3!). W kolejnej serii badaniu poddano telefon komórkowy [T], który był oceniany pod względem 4 kryteriów: czasu pracy (bate-

ria), ciężaru, wielkości wyświetlacza, oraz ekranu dotykowego. Ze względu na losowość, zastosowano 24 różne wersje kwestionariusza (4!). W ostatniej, trzeciej serii eksperymentu zbadano pięć obiektów: telefon komórkowy [T], mieszkanie [M], komunikację miejską [K], drukarkę [D] oraz laptop [L]. Każdy z nich był charakteryzowany przez 8 kryteriów, z tym, że liczba kryteriów dla każdego z obiektów była różna (ułożona w sposób losowy), co oznacza, że jeśli ankietowany oceniał np. [T] ze względu na 5 kryteriów, to [M] ze względu na 8 kryteriów, [K] ze względu na 7 kryteriów, a [L] ze względu na 6 kryteriów. Każdy z ankietowanych został poproszony o ocenę czterech z pięciu obiektów (T, M, K, D, L). Odpowiedzi były niezgodne, gdy współczynnik CR przekraczał wartość 0,10 (10%). Wyniki tych badań zostały zaprezentowane w Tabeli 3. Pokazują one, że bardzo wysoka frakcja niezgodnych odpowiedzi, ponad 50%, ma miejsce już przy analizie 3 kryteriów, a dla więcej niż 4 kryteriów odsetek ten przekracza 80%. W przypadku wszystkich badań, charakterystyczne było także występowanie współczynnika CR o wartości większej niż 0,50, a więc znacznie przekraczającej dopuszczalny poziom 0,10.

Tabela 3. Wyniki badań współczynnika niezgodności
Table 3. Results of the research on Consistency Ratio

Liczba kryteriów	Liczba poprawnie wypełnionych ankiet	Liczba ankiet, dla których $CR > 0,1$	Frakcja ankiet, dla których $CR > 0,1$
3	215	120	55,81%
4	162	129	79,63%
5	107	73	68,22%
6	106	88	83,02%
7	108	90	83,33%
8	102	87	85,29%

Źródło: obliczenia własne

Z powodu otrzymania wysokiego odsetka niezgodnych (nielogicznych) odpowiedzi, zdecydowano się na przeprowadzenie kolejnego eksperymentu polegającego na zbadaniu symetrii udzielanych odpowiedzi. Jak wcześniej wspomniano, zachowanie symetrii odpowiedzi jest warunkiem koniecznym (lecz niewystarczającym), dla zachowania zgodności osądów.

4.2. Badanie symetrii

Wskazania leżące powyżej wartości środkowej oznaczono znakiem „+”, natomiast leżące poniżej znakiem „-”. Wartości środkowe „obydwa kryteria mają takie same znaczenie” są pomijane. Przyjęto następujące oznaczenia:

$n_{(+)}$ - liczba wskazań leżących powyżej wartości środkowej,

$n_{(-)}$ - liczba wskazań leżących poniżej wartości środkowej,

$n_{(=)}$ - liczba wskazań oznaczających równoważność wskazań (takie samo znaczenie),

$N = n_{(+)} + n_{(-)} + n_{(=)}$ - liczba wszystkich porównań.

W celu dostosowania symboli do zapisu stosowanego w książkach statystycznych przyjęto:

$$n = n_{(+)} + n_{(-)}$$

$$m = n_{(+)}$$

Połowa wskazań (w sensie probabilistycznym) posiadać znak ”-”, a druga połowa znak „+”, czyli:

$$\frac{m}{n} \approx \frac{1}{2} \quad (8)$$

Zostanie zbadana hipoteza zerowa:

$$H_0: \frac{m}{n} = \frac{1}{2} \quad (9)$$

Wobec hipotezy alternatywnej:

$$H_1: \frac{m}{n} \neq \frac{1}{2} \quad (10)$$

Do weryfikacji hipotezy zerowej służy statystyka:

$$t_{obl} = \frac{\left| \frac{m}{n} - \frac{1}{2} \right|}{\sqrt{\frac{\frac{m}{n} \left(1 - \frac{m}{n} \right)}{n - 1}}} \quad (11)$$

Statystyka t ma przy założeniu prawdziwości hipotezy zerowej rozkład t studenta o $n-1$ stopniach swobody (można przyjąć, że dla dużych prób statystyka ma rozkład normalny).

Jeśli:

$$\left| t_{obl} \right| < t_{\alpha/2}, \quad (12)$$

to nie ma podstaw do odrzucenia hipotezy zerowej⁵. W przypadku przeciwnym $\left| t_{obl} \right| \geq t_{\alpha/2}$ hipotezę zerową należy odrzucić. Oznacza to, że kwestionariusz ankiety nie został właściwie zaprojektowany. W celu weryfikacji jakości zastosowanego kwestiona-

⁵ Stwierdzenie „nie ma podstaw do odrzucenia hipotezy zerowej” nie upoważnia do kategorycznego stwierdzenia, że kwestionariusz ankiety został zaprojektowany we właściwy sposób. Wskazuje jednak, że nie ma przeszkód, aby badany kwestionariusz był wykorzystywany do dalszych badań.

riusza o dwubiegunowej skali werbalnej z punktem środkowym w układzie pionowym poddano ponownej analizie zebrane dane na podstawie których został obliczony współczynnik zgodności (seria 3).

Tabela 4. Wyniki eksperymentu
Table 4. Results of the experiment

Seria / Obiekty	Kryteriów	Ankiety	Możliwych wersji	Porównań	$n_{(+)}$	$n_{(=)}$	$n_{(-)}$	$n_{(+)} + n_{(-)}$	t_{obl}	$t_{\alpha/2}$
1 – [S]	3	180	6	540	208	103	229	437	1,005	1,965
2 – [T]	4	96	24	576	252	87	237	489	0,678	1,965
3 – [T, M, K, D, L]	5	99	120	990	436	133	421	857	0,512	1,963
	6	98	720	1470	618	245	607	1225	0,314	1,962
	7	98	5040	2058	942	328	788	1730	3,716	1,961
	8	100	40320	2800	1127	465	1208	2335	1,677	1,961

Źródło: obliczenia własne

Dodatkowo zostały przeprowadzone dwie serie (1, 2) eksperymentu, w których wzięło udział od 96 do 180 osób (studentów), dokonując oceny dwóch obiektów: sklepu spożywczego [S] ze względu na 3 kryteria: szybkość obsługi, łatwość dojazdu i wybór produktów, oraz telefonu komórkowego [T] ze względu na 4 kryteria: czas pracy, ciężar, wielkość wyświetlacza oraz ekran dotykowy. Wyniki zostały zaprezentowane w Tabeli 4.

Jak wynika z Tabeli 4, w każdym przypadku oprócz 7 kryteriów uzyskano wynik $|t_{obl}| < t_{\alpha/2}$, wskazujący, że nie ma podstaw do odrzucenia hipotezy zerowej. To – upraszczając – wskazuje, że można uznać, że rozkład odpowiedzi jest symetryczny, a zatem ankieta została zaprojektowana w sposób właściwy. Jedynie podczas testowania hipotezy zerowej dla 7 kryteriów uzyskano wynik wskazujący na brak symetrii. Biorąc jednak pod uwagę wyniki dla pozostałej liczby kryteriów można stwierdzić, że jest to spowodowane błędem I rodzaju.

5. WNIOSKI

Odpowiedni kwestionariusz ankiety służący wspomaganie procesu pozyskiwania informacji pierwotnych ma fundamentalne znaczenie dla końcowych wyników analiz. Jeśli graficzna forma kwestionariusza jest niewłaściwa, to wtedy dane pierwotne mogą być obciążone błędami nielosowymi (systematycznymi). Przetwarzanie takich liczb zwykle prowadzi do nieprawdziwych wyników i – w konsekwencji – wniosków końcowych. Przeprowadzony eksperyment (test symetrii) wykazał, że kwestionariusz AHP został zaprojektowany we właściwy sposób, a mimo tego współczynnik CR był bardzo wysoki. Taki wynik sugeruje, że przyczyn niezgodności należy szukać gdzie indziej i potwierdza założenie, że zachowanie symetrii odpowiedzi nie jest warunkiem wystarczającym do zachowania logiczności osądów.

Przedstawiony sposób weryfikacji kwestionariusza ankiety dla skali dwubiegunowej nie dotyczy tylko metody AHP/ANP. Można go bowiem zastosować do każdej analizy, która korzysta z kwestionariusza dla skali symetrycznej, najczęściej dwubiegunowej. Zaniechanie sprawdzenia „jakości” kwestionariusza może prowadzić do sytuacji, w której uzyskane wyniki będą obciążone błędami. Wydaje się, że zaproponowana procedura weryfikacji kwestionariusza powinna być stosowana zawsze wtedy, gdy badacz wykorzystuje skalę dwubiegunową.

Autorzy pracy w kolejnym etapie badań zamierzają w podobny sposób sprawdzić jakość, przedstawionych w tym artykule, innych często stosowanych form graficznych kwestionariusza.

LITERATURA

- [1] Alonso J.A. i Lamata M.T.: *A Statistical Criterion of Consistency in the Analytic Hierarchy Process*, 2005. w:] Torra, V. i in. (Wyd.), MDAI 2005, 67-76, Springer-Verlag Berlin Heidelberg.
- [2] Babbie E.: *Badania społeczne w praktyce*, Wydawnictwo Naukowe PWN, Warszawa 2004.
- [3] Bradburn N., Sudman S. i Wansink B.: *Asking questions: the definitive guide to questionnaire design*, Jossey-Bass, San Francisco 2004.
- [4] Costa J.F., Wanderley A.J.M. i Cosenza C.A.N.: *A proposition to solve inconsistency problem in decision matrices using genetic algorithms*, “3rd International Conference on Production Research – Americas’ Region 2006” (ICPR-AM06).
- [5] Dillman D.: *Constructing the questionnaire. Mail and internet surveys*, John Wiley & Sons, New York 2000.
- [6] Fanning E.: *Formatting a Paper-based Survey Questionnaire: Best Practices.*, “Practical Assessment Research & Evaluation”, 10(12), 2005, s. 1-14.
- [7] Ishizaka H. i Labib A.: *Analytic Hierarchy Process and Expert Choice: Benefits and Limitations*, “ORInsight”, 22(4), 2009, s. 201-220.
- [8] Joshi R., Banwet D.K. i Shankar R.: *A Delphi-AHP-TOPSIS based benchmarking framework for performance improvement of a cold chain*, “Expert Systems with Applications”, 38(8), 2011, s. 10170-10182.
- [9] Goryl A., Jędrzejczyk Z., Kukuła K., Osiewalski J. i Walkosz A.: *Wprowadzenie do ekonometrii w przykładach i zadaniach*, pod red. Kukuły K., Wydawnictwo Naukowe PWN, Warszawa 2007.
- [10] Lee S. i Walsh P.: *SWOT and AHP hybrid model for sport marketing outsourcing using a case of intercollegiate sport*, “Sport Management Review”, 14(4), 2011, s. 361-369.
- [11] Miller D.: *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*, “The Psychological Review”, 63, 1956, s. 81-97.
- [12] Prusak A. i Stefanów P.: *Badania nad właściwościami operacyjnymi metody AHP*, „Folia Oeconomica Cracoviensia”, 52, 2011, s. 87-104.
- [13] Redline C. i Dillman D.: *The influence of alternative visual designs on respondents’ performance with branching instructions in self-administered questionnaires*, [w:] Groves R., Dillman D., Eltinge, J. i Little R. (Wyd.), “Survey Nonresponse”, Wiley, New York 2002.
- [14] Saaty T.L. i Ozdemir M.: *Why the magic number seven plus or minus two*, “Mathematical and Computer Modelling”, 38, 2003, s. 233-244.
- [15] Saaty T.L. i Peniwati K.: *Group Decision Making: Drawing out and Reconciling Differences*, RWS Publications, Pittsburgh, USA 2008.

- [16] Saaty T.L. i Vargas L.G.: *The Logic of Priorities*, Kluwer Nijhoff Publishing, Massachusetts 1982.
- [17] Saaty T.L.: *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, RWS Publications, Pittsburgh 1994.
- [18] Saaty T.L.: *Decision Making for Leaders: The Analytical Hierarchy Process for Decisions in a Complex World*, "The Analytical Hierarchy Process Series", 2, 1996, s. 71-74.
- [19] Saaty T.L.: *Decision making with the analytic hierarchy process*, "Int. J. Services Sciences", 1(1), 2008, s. 83-98.
- [20] [Sinuany-Stern et al. 2000] Sinuany-Stern Z., Mehrez A. i Hadad Y.: *An AHP/DEA methodology for ranking decision making units*, "International Transactions in Operational Research", 7(2), 2000, s. 109-124.
- [21] Sztumski J., *Wstęp do metod i technik badań społecznych*, „Śląsk”, Sp. z o.o., Katowice 1995.
- [22] Weathers D., Sharmab S. i Niedricha R.W.: *The impact of the number of scale points, dispositional factors, and the status quo decision heuristic on scale reliability and response accuracy*, "Journal of Business Research", 58, 2005, s. 1516-1524.
- [23] Webber S.A., Apostolou B. & Hassel J.M.: *The sensitivity of the analytic hierarchy process to alternative scale and cue presentations*, "European Journal of Operational Research", 96, 1996, s. 351-362.

GRAPHIC FORM OF QUESTIONNAIRE IN AHP/ANP RESEARCH

Analytic Hierarchy Process (AHP) and its extension, Analytic Network Process (ANP) are the most popular multi-criteria decision support methods. Data collection is based on pairwise comparisons of the elements of hierarchical or network models and must be made by the experts. The objective of this paper is to verify a thesis that high Consistency Ratio (CR) in the AHP/ANP methods, which indicates inconsistency (lack of logic) of the results, stems from the lack of symmetry of judgments in the wrongly designed questionnaire. The experiments were designed to test whether fraction of responses below and above the mid-point is probabilistically the same, and if not, whether it influences the value of CR. The hypothesis states that the fraction (quotient of responses below and above the scale mid-point) should be approximately $\frac{1}{2}$. The series test was used to verify this hypothesis. The research consisted of several experiments involving 8434 comparisons and over 400 respondents asked to evaluate different objects. The order and criteria of evaluation were different and random. The experiments used a 9-point, bipolar verbal and vertical scale in tabular form. In almost each case the results indicated symmetry of the responses. Hence, it can be said that the questionnaire was designed properly, and the source of high inconsistency is somewhere else. The proposed way of the questionnaire verification can be applied also in other methods. The results may also indicate that failure in controlling the „quality” of the questionnaire leads to the biases of final results.

Keywords: analytic hierarchy process, analytic network process, AHP, ANP, consistency, questionnaire

Badania zaprezentowane w artykule zostały sfinansowane ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji nr. DEC-2011/01/D/HS4/04006.

DOI: 10.7862/rz.2013.mmr.53

Tekst złożono w redakcji: kwiecień 2013

Przyjęto do druku: grudzień 2013