

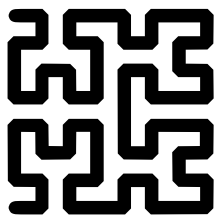


*Test Umiejętności Rozumowania
Dedukcyjnego (TRD) –
raport z konstrukcji narzędzia*

Raport Badawczy

numer: 6(6)/2016; opublikowany: 2 grudnia 2016.

KATARZYNA PALUSZKIEWICZ



REASONING
RESEARCH
GROUP

Badanie jest częścią projektu *Modelowanie rozumowań abdukcyjnych* finansowanego ze środków Narodowego Centrum Nauki (nr 2013/10/E/HS1/00172).

Spis treści

1	Wprowadzenie	2
2	Konstrukcja pozycji testowych	3
2.1	Pseudosłowa	3
2.2	Charakterystyka pozycji	3
2.3	Forma wniosku	4
3	Badanie	4
3.1	Różnice pomiędzy grupami	5
4	Własności psychometryczne	5
4.1	Wskaźnik trudności i mocy dyskryminacyjnej	5
4.2	Rzetelność	7
5	Analiza czynnikowa	7
6	Podsumowanie	8

Spis schematów

1	Przykładowe pozycje testowe	4
2	Wykres osypiska – czerwoną linią zaznaczono kryterium Kaisera	7

Spis tabel

1	Schematy pozycji testowych	4
2	Wyniki dla poszczególnych grup	5
3	Wskaźnik trudności dla poszczególnych pozycji	6
4	Wskaźniki trudności w grupach	6
5	Wskaźnik mocy dyskryminacyjnej dla poszczególnych pozycji	6
6	Ładunki czynnikowe dla poszczególnych czynników	8

1 Wprowadzenie

Jedną z głównych metod stosowaną w badaniach rozumowań dedukcyjnych są wnioskowania sylogistyczne. W literaturze psychologicznej określenie „sylogizm” jest stosowane dla szerszej klasy rozumowań niż przewiduje to klasyczna sylogistyka. Wyróżnia się najczęściej trzy rodzaje sylogizmów — liniowe, kategorioryczne oraz warunkowe (Matczak, 1994, s. 28). W poniższym raporcie określenie „sylogizm” będzie stosowane wyłącznie do schematów rozumowań składających się z dwóch przesłanek oraz wniosku, których składnikami są klasyczne zdania kategorioryczne:

- Każde S jest P — SaP
- Niektóre S są P — SiP
- Żadne S nie jest P — SeP
- Niektóre S nie są P — SoP

Innego rodzaju schematy wnioskowań będą nazywane po prostu wnioskowaniami o określonym charakterze. *Test Umiejętności Rozumowania Dedukcyjnego* składa się z 6 pozycji, każda z nich ma postać wnioskowania składającego się z dwóch przesłanek oraz wniosku. W skład narzędzia wchodzi zarówno sylogizmy, jak i wnioskowania relacyjne (o relacjach zachodzących pomiędzy obiektami).

Jednym z często spotykanych problemów w badaniach rozumowań dedukcyjnych jest występowanie systematycznych odchyżeń wyników (*biases*). Występowanie części z nich jest ściśle związane z treścią zdań używanych w badaniach. Badani chętniej uznają za poprawnie wyprowadzone wnioski: prawdziwe (w odniesieniu do wiedzy o świecie); takie, które wydają im się znane i prawdopodobne lub zgodne z własnymi przekonaniem (por. Evans, Barston i Pollard, 1983). Jednym ze sposobów wyeliminowania takich odchyżeń wyników jest używanie zmiennych literowych zamiast konkretnych nazw. Taki zabieg może jednak skłaniać badanych do samodzielnego podstawiania za zmienne literowe dowolnych nazw, co także może zaburzać wyniki. W teście TRD zastosowano odmienne rozwiązanie — zmienne literowe występujące w schematach zastąpiono materiałem abstrakcyjnych w postaci pseudosłów (ciągów liter danego języka możliwych do wymówienia, ale pozbawionych znaczenia). Podkreślony zostaje także fakt, że relacje opisane w przesłankach i wnioskach zachodzą pomiędzy zbiorami konkretnych obiektów (nawet jeżeli niedokładnie wiadomo, czym są te obiekty). Ocena poprawności wnioskowania sylogistycznego opartego na materiale abstrakcyjnym pozwala nam badać wyłącznie zdolność rozumowania dedukcyjnego i nie podlega wpływowi wiedzy czy przekonań badanych.

2 Konstrukcja pozycji testowych

2.1 Pseudosłowa

Pseudosłowa to ciągi liter danego języka, którego zachowują jego strukturę ortograficzną i fonologiczną. Przypominają one słowa danego języka, nie posiadają jednak znaczenia leksykalnego. Pseudosłowa użyte w teście TRD zostały zbudowane w oparciu o krótkie słowa języka polskiego (2–3 sylabowe), w których dokonano wymiany w obrębie samogłosek oraz spółgłosek. Na przykład, ze słowa „krowa” powstało pseudosłowo „growa”; z „mazur” – „muzor”. W konstrukcji zadań w teście TRD wykorzystano 18 pseudosłów:

- | | | | |
|------------|------------|------------|--------------|
| 1. arcich | 6. czaik | 11. growa | 16. rumburak |
| 2. bojuch | 7. czimar | 12. muzor | 17. tamno |
| 3. brukno | 8. gidziut | 13. oksek | 18. tykwa |
| 4. chowost | 9. girko | 14. parnał | |
| 5. ciamba | 10. glara | 15. pura | |

2.2 Charakterystyka pozycji

Test składa się z 6 pozycji. Trzy z nich mają postać klasycznych sylogizmów (pozycje 1–3), które składają się z klasycznych zdań kategorycznych. Pozycje sylogistyczne dobrano tak, aby uniknąć problemu z interpretacją wyrażenia *niektóre* – większość zdań użytych w konstrukcji to zdania ogólne. W przypadku pozycji 2, której wniosek jest zdaniem szczegółowotwierdzącym zdecydowano się na formę *jakieś C jest A*, aby wyeliminować ewentualne problemy z interpretacją wyrażenia *niektóre*. Trzy pozostałe (pozycje 4–6) natomiast zawierają zdania opisujące relacje większości lub mniejszości pomiędzy obiektami.

Większość zadań została zaczerpnięta z narzędzia „Test inteligencji płynnej. Rozumowanie dedukcyjne oparte na sylogizmach” (Grudzińska i Paluszkiewicz, 2011). Trzy pozycje testu zostały wykorzystane w formie niezmienionej (pozycje o numerach 2, 4 oraz 6), zmodyfikowano jedynie formę wyrażenia wniosku (zmiana zdania twierdzącego na pytanie rozstrzygnięcia). W przypadku pozycji 3, przesłanki pozostawiono bez zmian, natomiast wniosek przekształcono na zdanie ogólnotwierdzące. Pozostałe pozycje (1 oraz 5) mają formę analogiczną do pozostałych i zostały dobrane tak, by całe narzędzie zawierało trzy pozycje sylogistyczne oraz trzy pozycje relacyjne. Schematy wszystkich pozycji zostały zaprezentowane w Tabeli 1.

Następnym krokiem było zastąpienie zmiennych nazwowych (*A*, *B*, *C*) pseudosłowami. W przypadku pozycji sylogistycznych zastąpiono zmienne literowe (*a*, *e*, *i*) odpowiednimi wyrażeniami. W przypadku pozycji relacyjnych zdecydowano się na zróżnicowanie relacji, które będą opisywane. Założono, że sposób rozwiązywania zadania będzie taki sam (każda z tych relacji opisuje pewien wariant relacji większości-mniejszości). W pozycji 4 relacja,

Pozycja 1	Pozycja 2	Pozycja 3	Pozycja 4	Pozycja 5	Pozycja 6
AaB	AeB	AaB	$A < B$	$A < B$	$A < B$
BaC	BeC	BeC	$B < C$	$B < C$	$B > C$
AaC	CiA	CaA	$A < C$	$C < A$	$A > C$

Tabela 1: Schematy pozycji testowych

Każdy chowost jest arcichem.

Muzor jest krótszy niż czimar.

Każdy arcich jest rumburakiem.

Czimar jest krótszy niż oksek.

Czy każdy chowost jest rumburakiem?

Czy muzor jest krótszy niż oksek?

(a) pozycja sylogistyczne

(b) pozycja relacyjna

Schemat 1: Przykładowe pozycje testowe

w której pozostają ze sobą obiekty dotyczy długości (A jest krótsze niż B); w przypadku pozycji 5 dotyczy ona wysokości (A jest niższe od B , A jest wyższe od B); w pozycji 6 – wielkości (A jest mniejsze niż B , A jest większe niż B).

2.3 Forma wniosku

Zadaniem badanych była ocena poprawności wnioskowania zaprezentowanego w każdej z pozycji testowych. Aby uczynić ten proces bardziej naturalnym zdecydowano się na sformułowanie wniosków w formie pytań rozstrzygnięcia np. Czy każdy chowost jest rumburakiem?. Badani mogli odpowiedzieć na nie wybierając spośród trzech odpowiedzi: *na pewno tak*, *na pewno nie* oraz *nie można tego stwierdzić*. Założono, że taki format zaprezentowania zadania będzie dla badanych bardziej intuicyjny niż ocena prawdziwości zdania oznajmującego. Wygląd pozycji testowych zaprezentowano na Schemacie 1. W pytaniach wnioski przekształcono wyłącznie zdania twierdzące, aby uniknąć problemu z interpretacją odpowiedzi udzielanych na pytania zawierające negację.

3 Badanie

Test został wykorzystany w badaniu dotyczącym związków pomiędzy zdolnością rozwiązywania zadań dedukcyjnych a poziomem inteligencji ogólnej (Urbański, Paluszkiwicz i Urbańska, 2014). Poniżej opisane zostaną jedynie dane dotyczące Testu TRD.

Liczba osób badanych wyniosła 105 osób. Wszyscy badani byli studentami, średni wiek osób badanych to 21,69 ($SD = 1,44$). Średni wynik w całej grupie wyniósł 4,93 ($SD = 1,356$). Rozkład wyników ogólnych uzyskanych w teście nie jest zgodny z rozkładem normalnym ($W = 0,781$, $p < 0,001$). W całej grupie badanej wyróżniono trzy podgrupy: A ($N = 38$), B ($N = 40$) i C ($N = 27$). W grupach A i B znalazły się osoby tylko podstawowy kurs logiki, w grupie C – osoby po intensywnym kursie w zakresie logiki formalnej. Dodatkowo, w grupach B i C znajdowały się studenci kierunków, na których

Grupa	<i>M</i>	<i>SD</i>
A	3,68	1,416
B	5,53	0,679
C	5,81	0,396

Tabela 2: Wyniki dla poszczególnych grup

obowiązywała selektywna procedura przyjmowania na studia. Wyniki poszczególnych grup prezentuje Tabela 2. Rozkład wyników w żadnej z grup nie był zgodny z rozkładem normalnym (A: $W = 0,941$, $p = 0,045$; B: $W = 0,687$, $p < 0,001$; C: $W = 0,476$, $p < 0,001$).

3.1 Różnice pomiędzy grupami

Do przetestowania hipotezy o różnicy w poziomie poprawności rozwiązań w trzech grupach wykorzystano test Kruskalla-Wallisa. Wynik testu świadczy o występowaniu istotnych statystycznie różnic pomiędzy grupami ($\chi^2(2, 105) = 53,898$, $p < 0,001$). Analiza post-hoc wykazała występowanie istotnych różnic pomiędzy grupami A i B oraz grupami A i C (test Dunna, $p < 0,001$).

4 Własności psychometryczne

4.1 Wskaźnik trudności i mocy dyskryminacyjnej

Test TRD okazał się testem bardzo łatwym, jego średnia trudność w całej grupie badanych wyniosła 82,06%. Wskaźniki trudności dla poszczególnych pozycji przedstawia Tabela 3. Grupa badana składała się z trzech podgrup, w tym dwóch o poziomie inteligencji płynnej wyższej niż przeciętnej, dlatego przeanalizowano także trudności w poszczególnych grupach. Wyniki zostały zaprezentowane w Tabeli 4. Analiza wskaźników trudności w grupach B i C pokazuje, że test TRD praktycznie nie różnicuje badanych – średni wskaźnik trudności wynosi w obu przypadkach powyżej 90%. W przypadku osób o przeciętnej poziomie inteligencji (grupa A) test różnicuje badanych w stopniu zadowalającym; średnia trudność wynosi powyżej 60%.

Do obliczenia wskaźnika mocy dyskryminacyjnej wyróżniono dwie grupy po 28 osób – jedna z nich to osoby z najwyższymi wynikami, druga z najniższymi. Wskaźniki dla poszczególnych pozycji zostały zaprezentowane w Tabeli 5. Dwie pozycje charakteryzują się niższym od zalecanego (0,4) wskaźnikiem mocy dyskryminacyjnej. Są to pozycje 1 i 5 – dwie najlepiej wykonywane pozycje w grupie A (kolejno, trudność: 73,68% oraz 71,05%). Aby poprawić ich moc dyskryminacyjną należałoby wprowadzić w nich niewielkie zmiany np. zamienić kolejność zmiennych nazwowych we wnioskach, aby skomplikować proces rozwiązywania.

Pozycja	Wskaźnik trudności
1	88,57%
2	67,62%
3	83,81%
4	85,71%
5	87,62%
6	79,05%
średnia	82,06%

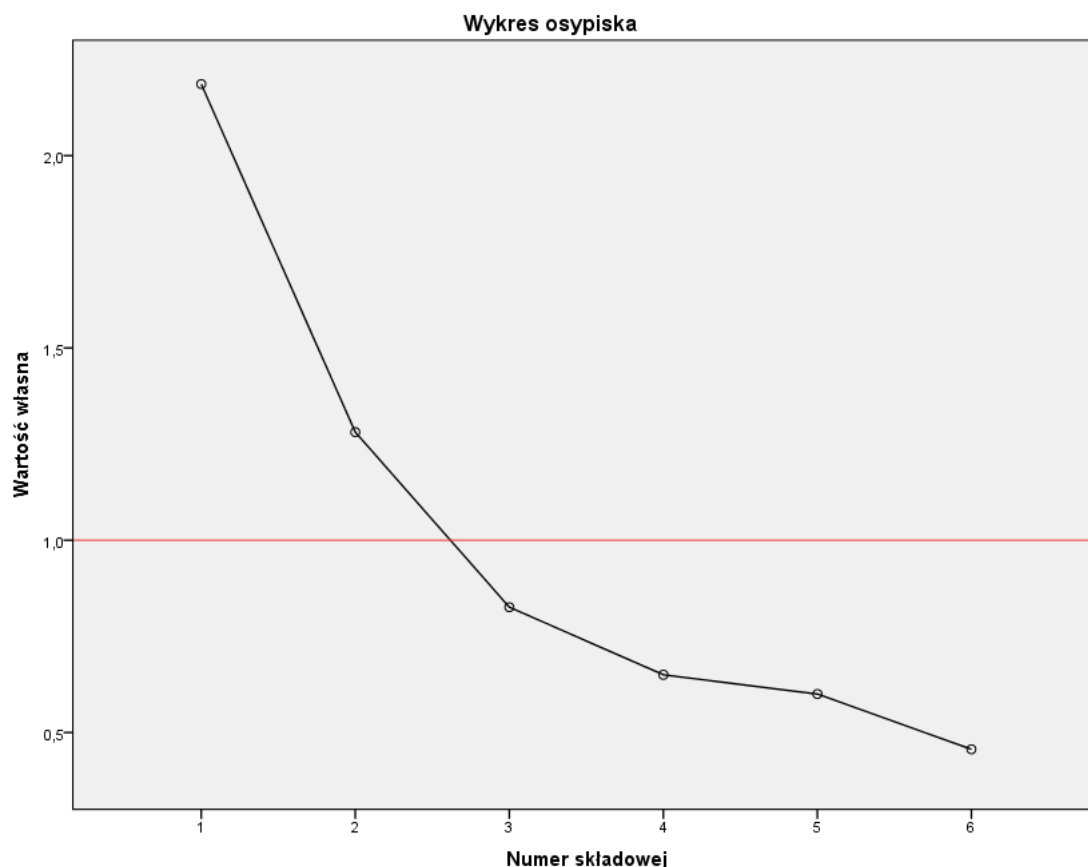
Tabela 3: Wskaźnik trudności dla poszczególnych pozycji

Pozycja	A	B	C
1	73,68%	97,50%	96,30%
2	39,47%	75,00%	96,30%
3	68,42%	95,00%	88,89%
4	60,53%	100,00%	100,00%
5	71,05%	95,00%	100,00%
6	55,26%	87,50%	100,00%
średnia	61,40%	91,67%	96,91%

Tabela 4: Wskaźniki trudności w grupach

Pozycja	Wskaźnik mocy dyskryminacyjnej
1	0,32
2	0,71
3	0,43
4	0,46
5	0,36
6	0,54

Tabela 5: Wskaźnik mocy dyskryminacyjnej dla poszczególnych pozycji



Schemat 2: Wykres osypiska – czerwoną linią zaznaczono kryterium Kaisera

4.2 Rzetelność

Do oceny rzetelności narzędzia użyto współczynnika λ_2 Guttmana, który uwzględnia trudność pozycji. Test TRD osiągnął rzetelność na poziomie 0,655.

5 Analiza czynnikowa

Aby sprawdzić, czy narzędzie ma strukturę czynnikową i dzieli się np. na podskale wyróżniane ze względu na typ pozycji wykonano eksploracyjną analizę czynnikową metodą głównych składowych (PCA; *principal component analysis*) z rotacją Varimax. Analiza wykresu osypiska (Schemat 2) sugeruje, że narzędzie nie ma struktury czynnikowej; kryterium Kaisera sugeruje dwa jako odpowiednią liczbę czynników. Ładunki czynnikowe dla poszczególnych pozycji i czynników zostały zaprezentowane w Tabeli 6. Pierwszy czynnik jest ładowany przez wszystkie pozycje, więc tak naprawdę narzędzie możemy uznać za spójne – nie ma potrzeby ani uzasadnienia dla wyróżniania podskal.

Pozycja	Czynnik 1	Czynnik 2
1	0,500	-0,620
2	0,561	0,419
3	0,665	
4	0,738	
5	0,458	0,711
6	0,653	-0,428

Tabela 6: Ładunki czynnikowe dla poszczególnych czynników

6 Podsumowanie

Test Umiejętności Rozumowania Dedukcyjnego okazał się testem o średnim poziomie rzetelności oraz niewielkiej trudności, zwłaszcza w grupach osób charakteryzujących się wysokim poziomem inteligencji ogólnej. Narzędzie ma strukturę jednoczynnikową, pomimo dwójakiego charakteru znajdujących się w nim pozycji.

Wyniki testu TRD nie różnicują badanych o inteligencji wyższej niż przeciętna. Wskaźnik średniej trudności w grupie A wskazuje natomiast, że test może być odpowiedni do różnicowania osób o przeciętnym poziomie inteligencji. Możliwe, że nadaje się także do badania osób młodszych, na przykład studentów początkowych lat oraz uczniów szkół ponadgimnazjalnych.

Bibliografia

- Evans, J. S. B., Barston, J. L., Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & cognition*, 11(3), 295–306.
- Grudzińska, J., Paluszkiewicz, K. (2011). *Test inteligencji płynnej. Rozumowanie dedukcyjne oparte na sylogizmach*. Niepublikowany raport z procesu projektowania narzędzia na zaliczenie przedmiotu teorie inteligencji. Poznań: Instytut Psychologii UAM.
- Maczak A. (1994). *Diagnoza intelektu*. Warszawa: Wydawnictwo Instytutu Psychologii.
- Urbański, M., Paluszkiewicz, K., Urbańska. (2014). *Deductive Reasoning and Learning: a Cross-Curricular Study*. Raport badawczy, Instytut Psychologii, Uniwersytet im. Adama Mickiewicza w Poznaniu.