

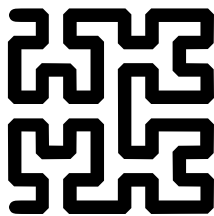


Giętkość dedukcyjna a inteligencja płynna oraz inne zdolności poznawcze

Raport Badawczy

numer: 4(4)/2016; opublikowany: 20 sierpnia 2016.

NATALIA ŻYLUK



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	3
1.1	Giętkość dedukcyjna	3
1.2	Giętkość dedukcyjna a inteligencja płynna	4
1.3	Giętkość dedukcyjna a inne zdolności poznawcze	5
1.3.1	Potrzeba poznawczego domknięcia	5
1.3.2	Styl epistemiczny	5
2	Metoda	7
2.1	Grupa badana	7
2.2	Narzędzia	7
2.2.1	<i>Test Matryc Progresywnych Ravena</i>	7
2.2.2	<i>Test Giętkości Dedukcyjnej</i>	8
2.2.3	<i>Skala Potrzeby Poznawczego Domknięcia</i>	9
2.2.4	<i>Standardized Epistemological Understanding Assessment</i>	10
2.3	Procedura badania	10
3	Wyniki	11
3.1	Statystyki opisowe	11
3.2	Analiza trudności	12
3.3	Analiza rzetelności	14
3.4	Wyniki analiz korelacyjnych	15
4	Dyskusja wyników	18
4.1	Analiza rzetelności	18
4.2	Analiza korelacyjna	18

1 WPROWADZENIE

Treścią niniejszego raportu jest opis badań nad związkami pomiędzy biegłością w zakresie trudnych rozumowań dedukcyjnych a inteligencją płynną oraz innymi charakterystykami poznawczymi jednostki (poziomu potrzeby poznawczego domknięcia, styl epistemiczny).

Wedle modelu inteligencji Catella-Horna-Carolla (modelu CHC) [Flanagan, 2008; Schneider i McGrew, 2012], umiejętność rozumowania dedukcyjnego – obok rozumowania indukcyjnego oraz ilościowego – uznawana jest za jedną z trzech zdolności składających się na komponent płynny inteligencji ogólnej. Bazując na owym modelu, jak również uwzględniając wyniki wcześniejszych badań nad związkami pomiędzy zdolnościami dedukcyjnymi a inteligencją płynną, jako główny cel analiz opisanych w niniejszym raporcie przyjęto sprawdzenie, czy istnieją statystycznie istotne związki pomiędzy zdolnościami poznawczymi, których wskaźnikiem są wyniki *Testu Matryc Progresywnych Ravena* [Raven, Raven i Court, 2011] a poziomem wykonania trudnych zadań dedukcyjnych, składających się na *Test Giętkości Dedukcyjnej* [Żyłuk, 2016]. Jako że rozwiązywanie testu wybranego do pomiaru umiejętności dedukcyjnych może dodatkowo angażować określone zdolności o charakterze metapoznawczym, do badania postanowiono włączyć pomiar wspomnianych dodatkowych zmiennych: poziomu potrzeby poznawczego domknięcia oraz posiadanego stylu epistemicznego.

1.1 GIĘTKOŚĆ DEDUKCYJNA

Giętkość dedukcyjna może być scharakteryzowana jako zdolność poznawcza manifestująca się, między innymi, w umiejętności wskazywania różnych zestawów przesłanek, z których wynikałby zadany wniosek [Żyłuk, 2016]. Nazwa badanej zmiennej ukuta została w oparciu o funkcjonujące w psychologii pojęcie giętkości poznawczej. Giętkość poznawczą najogólniej opisać można jako zdolność do przełączania pomiędzy myśleniem o odmiennych pojęciach oraz do myślenia o dwóch pojęciach jednocześnie [Scott, 1962]. Uznano, że sprawność w zakresie znajdowania zestawów przesłanek, z których logicznie wynikałby określony wniosek jest rodzajem specyficznie rozumianej „giętkości” – wymaga ona bowiem poznawczego przełączania między różnymi, potencjalnymi kandydatami na zestaw przesłanek oraz jednoczesnego myślenia o tych zestawach (lub inaczej: jednoczesnego przetwarzania informacji na ich temat). Zważywszy na charakter związku pomiędzy zdaniem, z którego wyprowadzany miałby być wniosek ustalony w zadaniu, przyjęto, że owa giętkość nazwana zostać może dedukcyjną. Choć wspomniana umiejętność wskazywania poprawnych zestawów przesłanek jest najbardziej bezpośrednim (najbardziej widocznym, a przez to najłatwiej podlegającym pomiarowi) przejawem zdolności, jaką jest giętkość dedukcyjna, nie zakłada się, że jest to jedyna owej zdolności manifestacja (stąd też „między innymi” na początku akapitu).

Z szerszej perspektywy, pomiar giętkości dedukcyjnej może być widziany jako pomiar specyficznego typu zdolności rozumowania dedukcyjnego. Biorąc pod uwagę rozróżnienie na łatwe i trudne rozumowania dedukcyjne [Urbański, Paluszkiwicz i Urbańska, 2014], test mierzący giętkość dedukcyjną – *Test Giętkości Dedukcyjnej* (opisany w części „Metoda”) należałoby uznać za test trudnych rozumowań dedukcyjnych. Dystynkcja „trudna”

versus „łatwa” dedukcja nie jest w żadnym wypadku ostra – bazuje ona bowiem na dość intuicyjnym założeniu, że pewne rodzaje zadań dedukcyjnych, z uwagi np. na większą liczbę procesów, jakie muszą zostać zaangażowane w celu poprawnego ich rozwiązania, sprawiają badanym więcej trudności aniżeli klasyczne, oparte na wykorzystaniu prostych sylogizmów, zadania dedukcyjne. Do testów mierzących proste rozumowania dedukcyjne zaliczyć można *Test Rozumowań Dedukcyjnych* (zbudowany przy wykorzystaniu właśnie klasycznych sylogizmów), natomiast do testów pozwalających na pomiar trudności w zakresie trudnej dedukcji należą *Test Rozumowań Erotetycznych* (angażujący rozumowania, w których pytania mogą pełnić rolę przesłanek i wniosków) czy też narzędzie *Style Rozumowania Dedukcyjnego* (złożone z zadań przybierających postać polisylogizmów) [Urbański, Paluszkiewicz i Urbańska, 2014]. Pojęcie „trudności” zadań dedukcyjnych może być rozumiane tu również zupełnie w dosłownym – ilościowym – sensie, albowiem okazuje się, że trudne zadania dedukcyjne są, w istocie, trudniejsze dla badanych aniżeli proste zadania dedukcyjne (o czym świadczą wskaźniki trudności uzyskiwane dla powyżej wymienionych narzędzi – por. tabela 1 na stronie 8 w przywoływanym powyżej raporcie). Zważywszy na fakt, że poprawne rozwiązywanie zadań w *Teście Giętkości Dedukcyjnej* wymaga nie tylko umiejętności rozpoznawania zachodzenia relacji wynikania logicznego, ale również zdolności kontroli własnych procesów poznawczych oraz „właściwego” zarządzania nimi w trakcie mierzenia się z problemem dedukcyjnym, można spodziewać się, że zadania tego rodzaju będą dla badanych trudniejsze aniżeli zadania oparte na prostych rozumowaniach sylogistycznych [Stenning i van Lambalgen, 2008].

1.2 GIĘTKOŚĆ DEDUKCYJNA A INTELIGENCJA PŁYNNA

Bazując na przywołanym uprzednio modelu inteligencji Catella-Horna-Carolla [Flanagan, 2008; Schneider i McGrew, 2012] należałoby się spodziewać, że umiejętność rozumowania dedukcyjnego, w jakimkolwiek jej aspekcie, wiązać powinna się z czynnikiem płynnym inteligencji ogólnej (a więc czynnikiem związanym ze zdolnościami do rozumowania, tworzenia pojęć oraz rozwiązywania problemów często wymagających wykorzystania nieznanymi informacjami czy procedur [Horn i Noll, 1997]). Jak jednak pozwalają wnosić wyniki badań zaprezentowane w przytaczanym powyżej raporcie Urbańskiego, Paluszkiewicz i Urbańskiej [2014], wskaźniki trudnych i łatwych rozumowań dedukcyjnych, choć ze sobą powiązane, charakteryzowane są przez odmienne wzorce powiązań ze wskaźnikiem inteligencji płynnej (w obrębie grup wyróżnionych ze względu na doświadczenie w zakresie nauki logiki), co może sugerować, że grupa zdolności dedukcyjnych jest grupą niejednorodną, i idąc dalej, że model CHC nie obejmuje pewnych aspektów tego typu umiejętności.

Badanie opisywane w niniejszym raporcie ukierunkowane było na eksploracyjne zbadanie powiązań pomiędzy wskaźnikiem zdolności do trudnej dedukcji a poziomem inteligencji płynnej, przy użyciu innej niż stosowanej przez przywoływanych autorów miary pierwszej z tych zmiennych.

1.3 GIĘTKOŚĆ DEDUKCYJNA A INNE ZDOLNOŚCI POZNAWCZE

1.3.1 POTRZEBA POZNAWCZEGO DOMKNIECIA

Zdefiniowane przez Kruglanskiego i Webstera [1996], pojęcie potrzeby domknięcia poznawczego odnosi się do dążenia jednostki do uzyskania jednoznacznej oraz pewnej odpowiedzi na pytania, z jakimi jest konfrontowana oraz do niechęci wobec wieloznaczności. Osoby cechujące się wysoką potrzebą poznawczego domknięcia mają skłonność do powierzchownej analizy napływających informacji, są przy tym w stanie szybko i sprawnie podejmować decyzje. Osoby o mniejszym poziomie tej cechy cechują się większą zdolnością do tolerowania niepewności i otwartością na nowe informacje. Niższy poziom potrzeby poznawczego domknięcia sprzyja także dokładnej i wnikliwej analizie sytuacji oraz ostrożności w zakresie formułowania ostatecznych sądów i poglądów na dany temat [Kossowska, Hanusz i Trejtowicz, 2012; Kruglanski i Webster, 1996].

Z uwagi na fakt, że rozwiązywanie wybranego testu rozumowań dedukcyjnych wymaga pogłębionej analizy zadań wchodzących w skład owego narzędzia; przełączania pomiędzy różnymi, potencjalnie prawidłowymi, zestawami przesłanek, z których miałyby wynikać zadane wnioski oraz zważywszy na fakt, że w kontekście niektórych zadań wchodzących w skład tego testu wymagane jest wskazanie więcej niż jednej prawidłowej odpowiedzi [Żyłuk, 2016], można zastanawiać się, czy wysoki poziom potrzeby poznawczego domknięcia byłby związany z niższymi wynikami testu mierzącego giętkość dedukcyjną. Nie przeprowadzono dotąd badań nad związkami pomiędzy potrzebą poznawczego domknięcia a poziomem wykonania zadań o charakterze zbliżonym do wybranego testu trudnych rozumowań dedukcyjnych, stąd też przyjrzenie się tego typu związkom (z – ponownie – czysto eksploracyjnej perspektywy) było jednym z celów badania opisywanego w niniejszym raporcie.

1.3.2 STYL EPISTEMICZNY

Styl epistemiczny może być scharakteryzowany jako posiadane przez jednostkę podejście w zakresie tego, czym jest wiedza, jaki ma charakter oraz w jaki sposób się ją zdobywa [Kuhn, Cheney i Weinstock, 2000]; innymi słowy, style epistemiczne są indywidualnymi koncepcjami na temat wiedzy i jej natury. Wedle przywoływanych Kuhn i współpracowników – autorów jednego z narzędzi do pomiaru posiadanego stylu epistemicznego – za rozwój poglądów na naturę wiedzy i przekonań odpowiadają zmiany relacji pomiędzy wymiarami obiektywizm-subiektywizm. Formowanie się „dojrzałego” stylu epistemicznego opisywane jest w literaturze jako proces mający swój początek w radykalnym obiektywizmie, następnie proces ten wiedzie przez subiektywizm, a kończy się na zintegrowaniu obu wymiarów.

Kuhn i współpracownicy wyróżnili [2000] trzy główne style epistemiczne: absolutysta, multiplista oraz ewaluatywista, które odpowiadają (odpowiednio) etapom dominacji obiektywizmu, dominacji subiektywizmu oraz etapu, na którym oba wymiary są skoordynowane. Co godne podkreślenia, przywoływani autorzy nie posługują się określeniem „stylu epistemicznego”, a „poziom rozumienia epistemologicznego” (*level of epistemological understanding*).

Absolutyści spostrzegają wiedzę jako byt obiektywny, całkowicie poznawalny i intelektualnie osiągalny. Według absolutystów, wszelka wiedza odnosi się rzeczywistości zewnętrznej względem podmiotu, która to rzeczywistość jest bezpośrednio poznawczo dostępna poznającemu. Dla przejścia z poziomu absolutysty na poziom multiplisty konieczne jest uświadomienie sobie niepewności wiedzy i jej subiektywnej strony. Wedle multiplistów źródło wiedzy związane jest z podmiotem, a nie rzeczywistością obiektywną. Uważają oni, że wszelkie stwierdzenia są jedynie opiniami, stąd mogą one cechować się jednakową słusnością. Według multiplisty, rzeczywistość nie jest bezpośrednio poznawalna i jest właściwą instancją, do której możemy się odwołać, chcąc ocenić odmienne racje. Styl ewaluatywisty charakteryzuje się dopuszczaniem jednoczesnej słusności pewnych poglądów (a więc jest w tym zakresie podobny do multiplisty), lecz z uwzględnieniem ich wartościowania (wymiar związany z obiektywizmem). Według ewaluatywisty pewne stwierdzenia mogą być zarazem słuszne, lecz jedno z nich (ze względu na ewidencję empiryczną bądź poparcie przekonującą argumentacją) może charakteryzować się większą słusnością.

Styl epistemiczny jest charakterystyką poznawczą jednostki silnie uzależnioną od domeny; Kuhn i współpracownicy ustalili, że styl epistemiczny może ujawniać się w kontekście 5 kategorii sądów: związanych z gustem, dotyczących oceny estetycznej, dotyczących wartości moralnych oraz dotyczących prawdy. Sądy dotyczące prawdy podzielone zostały na sądy dotyczące prawdy w kontekście społecznym oraz sądy dotyczące prawdy w kontekście świata fizycznego.

Nie przeprowadzono dotąd badań nad związkami pomiędzy posiadanymi stylami epistemicznymi a umiejętnościami rozumowania czy poziomem inteligencji. Z uwagi na fakt, że – wedle Kuhn i współpracowników – poziom poznawczego funkcjonowania jednostki jest silnie zdeterminowany jej poglądami na naturę wiedzy i przekonaniach, interesujące byłoby przyjrzenie się tego typu powiązaniom. Zapoznając się z charakterystyką trzech głównych stylów epistemicznych można odnieść wrażenie, że każdy z nich reprezentuje zupełnie inny sposób przetwarzania i oceny informacji w obrębie danej kategorii sądów; mierząc się z grupami odmiennych zdań na dany temat, osoby różniące się stylem epistemicznym aplikują inne kryteria ewaluacji poglądów wyrażanych w owych zdaniach. Zastanawiające, czy rozwiązywanie zadań abstrakcyjnych, oderwanych od domeny mogłoby wiązać się w jakiś sposób z posiadanym stylem epistemicznym i – jeżeli tak – stylem w obrębie którego typu poglądów. Co godne odnotowania, istnieją badania pozwalające podejrzewać, że posiadany styl epistemiczny mógłby wiązać się także z poziomem potrzeby poznawczego domknięcia [Peter, Rosman, Mayer, Leichner i Krampen, 2015]. W badaniu Peter i współpracowników [2015] stwierdzono, że „wyrafinowanie epistemiczne” (*epistemic sophistication*), którego przejawem jest m.in. przejawianie poglądów ewaluatywistycznych, było negatywnie skorelowane z poziomem potrzeby domknięcia poznawczego: im bardziej wyrafinowane epistemiczne poglądy cechowały jednostkę, tym mniej poznawczo była ona domknięta. Należy przy tym zaznaczyć, że w opisywanym badaniu wykorzystano inne miary poziomu rozumienia epistemologicznego niż w trakcie analiz będących przedmiotem opisu niniejszego raportu.

Biorąc pod uwagę luki w zakresie badań nad związkami pomiędzy posiadanym stylem epistemicznym a innymi aspektami poznawczego funkcjonowania jednostki oraz pewne wstępne podejrzenia dotyczące możliwych powiązań pomiędzy badanymi zmiennymi, do opisywanego badania postanowiono włączyć test *Standardized Epistemological Understanding Assessment*.

2 METODA

2.1 GRUPA BADANA

Grupę badanych stanowiło 47 studentów różnych kierunków w wieku od 20 do 25 lat ($M = 22,72$; $SD = 1,246$). Każda z osób badanych studiowała na co najmniej trzecim roku studiów (III rok – 23 osoby, IV rok – 12 osób, V rok – 11 osób, łącznie IV i V rok – 1 osoba). Wśród uczestników badania znalazło się 26 kobiet oraz 21 mężczyzn (nie odnotowano istotnych różnic w zakresie liczebności obu grup: $\chi^2 = 0,532$, $p > 0,05$). Najczęściej wskazywanymi kierunkami były: kognitywistyka ($n = 5$), automatyka i robotyka ($n = 4$), wzornictwo, psychologia, zarządzanie środowiskiem oraz biotechnologia (wszystkie kierunki po $n = 3$). Jeden z uczestników zadeklarował równoczesne studiowanie dwóch kierunków (kognitywistyki i psychologii, na, odpowiednio V i IV roku – por. wyżej).

W badaniu zastosowano dobór celowy z zamiarem uzyskania grupy badanych homogenicznej pod względem wieku oraz znajdujących się na podobnym etapie edukacji. Badanych rekrutowano poprzez bezpośrednie zapraszanie do udziału w badaniu oraz poprzez ogłoszenia zamieszczane w mediach społecznościowych. Część uczestników zgłosiła swoją chęć udziału w badaniu za sprawą rekomendacji osób, które wcześniej w nim uczestniczyły. Badani nie otrzymywali wynagrodzenia za udział w badaniu.

2.2 NARZĘDZIA

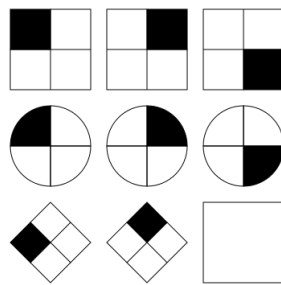
W badaniu wykorzystano 4 testy:

- *Test Matryc Progresywnych Ravena* (w wersji dla zaawansowanych) (*Raven's Advanced Progressive Matrices test*; APM) [Raven, Raven i Court, 2011],
- *Test Giętkości Dedukcyjnej* (*Deductive Flexibility Test*; DFT) [Żyłuk, 2016],
- *Skalę Potrzeby Poznawczego Domknięcia* w wersji skróconej (*Need for Cognitive Closure Scale*; NFCS) [Kossowska, Hanusz i Trejtowicz, 2012],
- test stylów epistemicznych – *Standardized Epistemological Understanding Assessment* (SEUA) w wersji papier-ołówek [Żyłuk i in., 2016].

W dalszej części raportu zamiast pełnych nazw narzędzi stosowane będą niekiedy angielskojęzyczne ich skrótowce. Zestawienie wskaźników rzetelności zastosowanych narzędzi znajduje się w części 3.3.

2.2.1 Test Matryc Progresywnych Ravena

Test Matryc Progresywnych Ravena jest narzędziem liczącym 48 pozycji testowych, prezentowanych w dwóch zestawach: Serii I, treningowej, na którą składa się 12 zadań oraz Serii II, będącej testem właściwym, złożonej z 36 zadań. Jedną pozycję testową stanowi macierz (3x3) zbudowana z bodźców obrazkowych o postaci figur, z jednym miejscem pustym (por. obrazek). Bodźce obrazkowe składające się na macierz ułożone są wedle ukrytej reguły. W przypadku każdej z pozycji testowych, zadaniem badanego jest odkrycie owej prawidłowości oraz wskazanie fragmentu, który prawidłowo – z uwagi na ową regułę – uzupełnia zaprezentowany wzór. Badany wybiera swoją odpowiedź spośród ośmiu zaproponowanych. Pozycje testowe ułożone są wedle rosnącego stopnia trudności. Na wynik ogólny składa się liczba poprawnie rozwiązanych zadań (zwykle tylko z Serii II, maksymalnie 36 punktów).



Rysunek 2.1: Przykładowa matryca przypominająca wykorzystywane w teście APM

Test Matryc Progresywnych Ravena uważany jest za dobrą miarę zdolności do rozumowania, jako jednego z komponentów składających się na czynnik płynny inteligencji ogólnej [Blair, 2010]. W badaniu opisywanym w niniejszym raporcie zastosowana została zaawansowana wersja testu, rekomendowana dla adolescentów z wyższym niż przeciętny poziomem inteligencji oraz dla dorosłych [Raven, Raven i Court, 2011]. *Test Matryc Progresywnych Ravena* w wersji dla zaawansowanych może być stosowany bez limitu czasowego, jak też z ograniczeniem czasu. W pierwszym przypadku użycie APM pozwala na ocenę sprawności intelektualnej (*efficiency*) jednostek, umożliwiając różnicowanie osób badanych w zakresie bardzo wysokich wyników tego testu. W warunkach z ograniczeniem czasowym, zastosowanie APM może posłużyć do oszacowania efektywności wysiłku intelektualnego [Raven, Raven i Court, 2011]. W opisywanym badaniu wykorzystana została wersja APM z ograniczeniem czasowym (do 3 minut na zapoznanie się – bez konieczności zaznaczania odpowiedzi – z serią I oraz dokładnie 40 minut na rozwiązanie Serii II). Do analiz wykorzystano wyniki uzyskane w Serii II.

2.2.2 Test Giętkości Dedukcyjnej

Test Giętkości Dedukcyjnej jest testem logicznym opracowanym przez Żyłuk i Urbańskiego [Żyłuk, 2016], złożonym z 8 zadań. Pojedynczą pozycję testową stanowi zestaw sześciu klasycznych zdań kategorycznych: jedno zdanie pod kreską (wniosek) oraz pięć zdań nad kreską (potencjalne przesłanki). Zadaniem osoby badanej jest wybór wszystkich minimalnych zestawów zdań znad kreski, z których wynikałby zadany wniosek (przy czym przez „zestaw” rozumie się także pojedyncze zdanie). W konstrukcji pozycji testowych wybranych wykorzystano pseudosłowa. Przykładowe zadanie testu prezentuje się następująco:

- (1) Niektóre czuły są mutkami.
- (2) Niektóre mutki są zymkami
- (3) Każda czuła jest mutką.
- (4) Każda czuła jest zymkiem.
- (5) Nie jest tak, że żaden zymek nie jest mutką.

Niektóre zymki są mutkami.

Rozwiązaniem zadania są następujące zestawy przesłanek: przesłanka nr 2, przesłanka nr 5, para przesłanek 3 i 4 oraz para przesłanek 1 i 4.

W skład testu weszły pozycje testowe, w przypadku których możliwe jest znalezienie od 0 do 5 prawidłowych zestawów zdań nad kreską. Co istotne, dane zdanie ze zbioru nad kreską może być „użyte” w różnych zestawach przesłanek w jednym zadaniu. Dla każdego zadania badany mógł otrzymać liczbę punktów z przedziału od 0 do 1. Szczegółowa charakterystyka konstrukcji pozycji testowych, wraz z wyczerpującym omówieniem budowy wskaźnika poprawności wykonania pojedynczego zadania, zostały zaprezentowane w raporcie z konstrukcji testu [Żyluk, 2016].

Pozycje testu cechują się zróżnicowaną trudnością (por. tabela 3.3.); w procesie konstrukcji narzędzia starano się, aby zadania różniły się od siebie pod względem poziomu trudności, ale aby zarazem ich średnia trudność oscylowała w okolicach 50% (zgodnie z rekomendacją Hornowskiej [2007]). Wynik ogólny w teście stanowi wartość średnią z liczby punktów uzyskanych we wszystkich zadaniach testu (wynik z przedziału 0–1).

DFT skonstruowany został jako test mierzący zdolność w zakresie rozwiązywania trudnych zadań dedukcyjnych, tym niemniej poprawność odpowiedzi badanych rozwiązujących ów test nie jest zdeterminowana wyłącznie umiejętnością rozpoznawania zachodzenia relacji wynikania, a zależy także od poziomu umiejętności poznawczych pozwalających na sprawne przełączanie się pomiędzy potencjalnymi, różnymi „kandydatami” na rozwiązania.

2.2.3 Skala Potrzeby Poznawczego Domknięcia

Do pomiaru poziomu potrzeby domknięcia poznawczego zastosowano *Skalę Potrzeby Poznawczego Domknięcia* w wersji skróconej. Skrócona wersja *Skali Potrzeby Poznawczego Domknięcia* jest narzędziem, na które składa się 15 pozycji testowych, mających formę zdań oznajmujących. Zadaniem badanego jest ustosunkowanie się do każdego ze stwierdzeń-pozycji poprzez zaznaczanie odpowiedzi na skali sześciostopniowej, gdzie: 1 oznacza „całkowicie się nie zgadzam”, natomiast 6 – „całkowicie się zgadzam”. Wyróżniono 5 podskal narzędzia: preferowanie porządku, preferowanie przewidywalności, nietolerancja wieloznaczności, zamkniętość umysłowa, zdecydowanie. Treść przykładowej pozycji brzmiała: „Unikam brania udziału w wydarzeniach, nie wiedząc, czego mogę się po nich spodziewać” (podskala preferencja przewidywalności).

Im wyższy wynik w teście NFCS otrzyma osoba, tym wyższy poziom poznawczego domknięcia ją cechuje. W skróconej wersji narzędzia uzyskać można od 15 do 90 punktów (od 3 do 18 w obrębie jednej podskali). W opisywanym badaniu jako zmienne traktowano wynik ogólny oraz wyniki pięciu podskal.

2.2.4 *Standardized Epistemological Understanding Assessment*

Narzędzie SEUA stanowi zmodyfikowaną i zaadaptowaną wersję testu opracowanego przez Kuhn, Weinstocka i Cheney'a [2000]. Test SEUA – jako nowa wersja narzędzia – stosowany był dotąd w wersji wywiadu z uczestnikiem [Żyłuk i in., 2016], natomiast w trakcie opisywanego badania po raz pierwszy zdecydowano się na wykorzystanie go w formie papier-ołówek.

Zastosowanie testu SEUA pozwala na ocenę posiadanego stylu epistemicznego tj. podejścia, jakie ma jednostka w zakresie tego, czym jest wiedza, jaki ma charakter oraz w jaki sposób się ją zdobywa. Autorzy oryginalnego narzędzia wyróżnili trzy główne style epistemiczne: absolutysta, multiplista oraz ewaluatywista [Kuhn, Cheney i Weinstock, 2000]. Pełen opis wszystkich trzech stylów znajduje się we wprowadzającej części raportu. Styl epistemiczny może ujawniać się w kontekście 5 kategorii sądów: związanych z gustem, dotyczących oceny estetycznej, dotyczących wartości moralnych, prawdy w kontekście społecznym oraz prawdy w kontekście świata fizycznego.

Test SEUA złożony jest 25 pozycji testowych, dotyczących wymienionych powyżej kategorii sądów (po 5 pozycji na jedną grupę sądów). Na pojedynczą pozycję testową składa się para stwierdzeń prezentujących poglądy dwóch osób. Przykładowa pozycja prezentuje się następująco:

Robert sądzi, że śluby powinny odbywać się przed południem.

Michał sądzi, że śluby powinny odbywać się po południu.

W kontekście każdej pozycji testu badany odpowiada na jedno lub dwa pytania: po pierwsze, o to, czy czy tylko jeden z tych poglądów może być słuszny, czy oba mogą być do pewnego stopnia słuszne; a następnie, jeżeli zaznaczył, że oba poglądy mogą być do pewnego stopnia słuszne, o to, czy jeden z tych poglądów może być bardziej słuszny od drugiego.

Wynik w teście wskazuje na poziomie zintegrowania wymiarów obiektywizm-subiektywizm w zakresie owych 5 kategorii sądów. Test SEUA pozwala na uzyskanie wyniku zarówno jakościowego (profil literowy złożony z liter A, M, E), jak i ilościowego (wartości liczbowe; w obrębie każdej podskali możliwe jest do uzyskania od 5 do 15 punktów). Analizie podlegać mogą wyłącznie wartości uzyskane w kontekście podskal narzędzia, odpowiadającym wspomnianym 5 domenom; wynik ogólny testu nie jest poddawany interpretacji.

2.3 PROCEDURA BADANIA

Badanie zostało przeprowadzone w okresie od maja do sierpnia 2016 roku. Badania odbywały się w laboratorium Reasoning Research Group (Instytut Psychologii Uniwersytetu im. Adama Mickiewicza w Poznaniu) bądź, w szczególnych przypadkach, poza laboratorium (np. w bibliotece), z zachowaniem identycznych warunków testowania. Uczestnicy badani byli indywidualnie lub w grupach (maksymalnie czteroosobowych).

Przed rozpoczęciem wypełniania testów uczestnicy zapoznawali byli z tematyką i celem badania, jak również kwestiami formalnymi dotyczącymi udziału w nim. Badanych informowano o anonimowości badania, opcji rezygnacji z uczestnictwa w nim w dowolnym

momencie bez podawania powodu oraz o możliwości zadawania pytań w czasie trwania spotkania. Badani zostali powiadomieni o sposobach uzyskania informacji o wynikach, jakie zdobędą w testach (kontakt z osobą kierującą projektem lub zapisanie adresu e-mail, na który osoba przeprowadzająca badanie miała wysłać wyniki). Wszelkie informacje na temat badania uczestnicy otrzymali również w wersji papierowej. Jako zgodę na udział w badaniu potraktowano wysłuchanie instrukcji i przystąpienie do uzupełniania pierwszego testu.

Na badanie składały się dwie, około pięćdziesięciminutowe, sesje, rozdzielone pięciminutową przerwą. W trakcie pierwszej części badani wypełniali test APM (40 minut) oraz NFCS (ok. 10 minut), następnie, po przerwie, DFT (ok. 25 minut) oraz SEUA (ok. 25 minut). *Test Matryc Progresywnych Ravena* był jedynym testem przeprowadzonym z zastosowaniem ścisłego ograniczenia czasowego. Przed rozpoczęciem wypełniania każdego z testów badany był zapoznawany ze szczegółową instrukcją jego rozwiązywania. Wszystkie testy stosowane były w języku polskim.

Ustalenie porządku testów w sposób powyżej opisany motywowane był kilkoma względami. Założono, że APM oraz DFT, jako testy zdolności złożone z wysoce abstrakcyjnych i angażujących poznawczo zadań, nie powinny być wypełniane zaraz po sobie. Ponadto, uznano za kluczowe, aby do uzupełniania APM i DFT badani podchodzili możliwie nieznużeni – tak, aby mogli wykazać się swoimi możliwościami w trakcie ich rozwiązywania (w przypadku APM – w ramach narzuconego limitu czasowego); z tego powodu badanie postanowiono rozpocząć od APM, a DFT ułożyć tuż po przerwie. Taka kolejność prezentacji narzędzi pozwalała również niejako wykorzystać zmęczenie mogące pojawić się po wypełnieniu APM i DFT: uzupełnianie NFCS i SEUA – testów postaw i przekonań – po rozwiązaniu (odpowiednio) APM i DFT, mogło sprzyjać udzielaniu bardziej intuicyjnych odpowiedzi w kontekście tych pierwszych, bez poprzedzenia ich zaznaczenia głębokim namysłem.

Ostatecznie zdecydowano o umieszczeniu NFCS tuż po APM, a SEUA po DFT, co pozwoliło na wydzielenie dwóch, zbliżonych pod względem długości trwania, bloków badania.

3 WYNIKI

Wszystkie z analiz opisanych w niniejszym rozdziale przeprowadzane były z wykorzystaniem oprogramowania IBM SPSS w wersji 23.0. Przyjęto poziom istotności $\alpha = 0,05$.

3.1 STATYSTYKI OPISOWE

Tabela 3.1 zawiera podstawowe statystyki opisowe analizowanych zmiennych oraz wyniki testu Kołmogorowa-Smirnowa (wraz z uzyskanymi poziomami istotności), użytego w celu zbadania zgodności rozkładów wyników poszczególnych testów z rozkładem normalnym. Jak sugeruje treść tabeli, rozkład zgodny z rozkładem normalnym przyjęły wyniki APM, DFT, jak również wynik ogólny NFCS oraz wyniki dwóch z pięciu podskal NFCS (podskala preferowanie porządku i zdecydowanie). Wyniki pozostałych trzech podskal NFC oraz wszystkich podskal narzędzia SEUA przyjęły rozkład istotnie różniący się od normalnego.

Tabela 3.1: Statystyki opisowe

test	n	M	SD	Me	Min	Max	Test K-S	
							Z	p
APM	47	26,34	4,425	25	14	34	0,114	0,157
DFT	47	0,55	0,189	0,55	0,177	0,975	0,063	0,200
NFCS: ogólny	47	52,32	6,985	52	29	63	0,115	0,154
NFCS: pref. przewid.	47	9,87	2,708	10	4	16	0,136	0,030
NFCS: pref. porz.	47	11,47	3,161	11	3	18	0,114	0,160
NFCS: nietol. wielozn.	47	13,53	2,977	14	4	18	0,222	<0,001
NFCS: zam. umysłowa	47	7,87	1,907	8	3	12	0,191	<0,001
NFCS: zdecydowanie	47	9,53	3,028	9	3	15	0,112	0,185
SEUA: gust	46	10,54	1,394	10	8	14	0,196	<0,001
SEUA: estetyka	46	10,59	1,147	10	10	14	0,413	<0,001
SEUA: wartości moralne	46	11,15	2,582	11	5	15	0,154	0,008
SEUA: prawda: świat społ.	46	13,09	2,009	13	5	15	0,200	<0,001
SEUA: prawda: świat fiz.	46	11,96	2,590	12,5	5	15	0,156	0,006

Co godne odnotowania, w przypadku narzędzia SEUA konieczne było wykluczenie z analiz jednego z arkuszy testowych; do analiz włączono zatem 46, a nie 47 wyników testu SEUA (jak sugeruje kolumna z licznosciami grup). Arkusz ów odrzucono ze względu na nieprawidłowe uzupełnienie go przez badanego. Sposób uzupełnienia arkusza usuniętego z analiz wskazywał na całkowity brak zrozumienia przez uczestnika badania instrukcji rozwiązywania SEUA.

3.2 ANALIZA TRUDNOŚCI

W przypadku *Testu Matryc Progresywnych Ravena* oraz *Testu Giętkości Dedukcyjnej* możliwe jest obliczenie wskaźników trudności. Miara, jaką jest wskaźnik trudności, ma sens interpretacyjny jedynie wówczas, gdy w odniesieniu do pytań wchodzących w skład danego narzędzia istnieje kategoria odpowiedzi prawidłowej [Hornowska, 2007]. W konsekwencji powyższego, wskaźnik ten oblicza się zwykle wyłącznie dla testów zdolności poznawczych.

W kontekście *Testu Matryc Progresywnych Ravena* wskaźnik trudności danej pozycji testowej obliczany był jako procent badanych, który wskazał prawidłową odpowiedź rozwiązując to zadanie. Nie zdecydowano się na obliczanie wskaźnika rozumianego jako stosunek liczby prawidłowych odpowiedzi do wszystkich udzielonych ze względu na trudności interpretacyjne dotyczące braku zaznaczenia odpowiedzi – nie było podstaw do rozstrzygnięcia, czy brak zaznaczenia odpowiedzi w przypadku danej pozycji testowej wynika z niewiedzy badanego (badany pominął, bo nie wiedział, co zaznaczyć; a takie zachowanie należałoby zapewne interpretować jako swojego rodzaju udzielenie odpowiedzi) czy jest konsekwencją ścisłego ograniczenia czasowego (badany nie zdążył w ogóle dojść do pytania bądź zostawił je „na później”, ale nie miał sposobności powrotu).

W tabeli 3.2 znalazły się wskaźniki trudności obliczone dla każdego z 36 zadań testu APM, jak również średni wskaźnik trudności z wszystkich uzyskanych wskaźników. Jak sugerują wartości obliczonej miary, w przypadku dwóch zadań wszyscy badani wskazali prawidłowe ich rozwiązania. Najtrudniejszym pytaniem okazało się być pytanie ostatnie. Warto podkreślić, że względu na to, że test uzupełniany był ze ścisłym ograniczeniem czasowym, w kontekście niektórych pytań niższy odsetek prawidłowych wskazań mógł być po części konsekwencją ich ułożenia.

Tabela 3.2: Wskaźniki trudności zadań APM

Zadanie	Wskaźnik trudności
1.	87,23%
2.	100%
3.	93,62%
4.	95,74%
5.	93,62%
6.	97,87%
7.	100%
8.	95,74%
9.	95,74%
10.	89,36%
11.	95,74%
12.	89,36%
13.	78,72%
14.	95,74%
15.	87,23%
16.	91,49%
17.	85,11%
18.	89,36%
19.	85,11%
20.	78,72%
21.	72,34%
22.	76,60%
23.	89,36%
24.	48,94%
25.	57,45%
26.	53,19%
27.	63,83%
28.	34,04%
29.	38,30%
30.	55,32%
31.	53,19%
32.	38,30%
33.	42,55%
34.	38,30%
35.	38,30%
36.	8,51%
średnia:	73,17%

Tabela 3.3 zestawia wskaźniki trudności obliczone dla wszystkich 8 zadań wchodzących w skład DFT oraz wskaźnik średni. Przypomnijmy, że w przypadku każdej pozycji testowej DFT badany mógł otrzymać wynik (wskaźnik poprawności) z przedziału 0–1. Wskaźnik trudności dla danego zadania obliczano jako średni wskaźnik jego poprawności w całej próbie, wyrażony procentowo.

Jak sugerują wskaźniki zaprezentowane w poniższej tabeli, stosunkowo prostymi pozycjami testowymi okazały się być pozycje 1, 6 i 7 (średni wskaźnik poprawności wynoszący około 75%). Zadaniem, które sprawiło badanym najwięcej trudności było zadanie nr 3 – średni wskaźnik poprawności wyniósł niecałe 9%. Warto odnotować, że pozycja testowa o numerze 3 jest wyjątkowa na tle pozostałych wchodzących w skład testu – prawidłowym rozwiązaniem w jej przypadku jest brak wskazania jakiegokolwiek zestawu przesłanek. Również w trakcie badania pilotażowego zadanie to okazało się być najtrudniejsze dla badanych [Żyłuk, 2016].

Tabela 3.3: Wskaźniki trudności zadań DFT

Zadanie	Wskaźnik trudności
1.	74,46%
2.	38,01%
3.	8,51%
4.	72,26%
5.	47,71%
6.	74,26%
7.	75,07%
8.	53,38%
średnia:	55,46%

3.3 ANALIZA RZETELNOŚCI

W celu oszacowania poziomu rzetelności narzędzi wybranych do badania posłużono się dwoma współczynnikami: α Cronbacha oraz λ_2 Guttmana. Miara α Cronbacha jest najpopularniejszym współczynnikiem pozwalającym na określenie homogeniczności pozycji wchodzących w skład danego narzędzia. Współczynnik λ_2 Guttmana jest wskaźnikiem nieco bardziej złożonym i zarazem mniej popularnym niż α Cronbacha; jego zastosowanie umożliwia oszacowanie rzetelności narzędzi złożonych z zadań o zróżnicowanej trudności [Guttman, 1945]. Interpretacja wielkości obu wskaźników jest analogiczna.

Z uwagi na fakt, że testy APM i DFT są narzędziami złożonymi z pozycji testowych o zróżnicowanej trudności, w celu obliczenia rzetelności obu testów posłużono się współczynnikiem λ_2 . Wskaźnik α Cronbacha obliczony został dla wyniku ogólnego i wyników podskal NFCS oraz dla podskal narzędzia SEUA. Uzyskane wskaźniki rzetelności zaprezentowano w tabeli 3.4.

Tabela 3.4: Wyniki analiz rzetelności

test	poziom rzetelności	
	α Cronbacha	λ_2 Guttmana
APM	-	0,789
DFT	-	0,820
NFCS: ogólny	0,626	-
NFCS: pref. przewid.	0,659	-
NFCS: pref. porz.	0,847	-
NFCS: nietol. wielozn.	0,711	-
NFCS: zam. umysłowa	0,552	-
NFCS: zdecydowanie	0,763	-
SEUA: gust	0,436	-
SEUA: estetyka	0,764	-
SEUA: wartości moralne	0,647	-
SEUA: prawda: świat społ.	0,721	-
SEUA: prawda: świat fiz.	0,670	-

3.4 WYNIKI ANALIZ KORELACYJNYCH

Jak podkreślano we wprowadzającej części raportu, opisywanie badanie miało charakter wyłącznie eksploracyjny. Analizy korelacyjne ukierunkowane na zbadanie związków pomiędzy zmierzonymi zmiennymi zostały przeprowadzone w trzech etapach. Etapy te wydzielono zważywszy na hierarchię ustalonych celów badawczych.

W trakcie pierwszego z etapów badane były związki pomiędzy wynikami APM a wynikami uzyskanymi w DFT. Jak nadmieniono we wprowadzeniu niniejszego raportu, powiązania między tymi miarami były głównym przedmiotem zainteresowania w trakcie opisywanego badania. Podczas drugiego etapu sprawdzano, czy wyniki DFT są powiązane z wynikami NFCS (zarówno z wynikiem ogólnym i z wynikami podskal) oraz czy wyniki DFT są skorelowane z wynikami podskal testu SEUA. W trakcie trzeciego etapu analiz przeprowadzono analizy dodatkowe celem sprawdzenia, czy pomiędzy zmiennymi uwzględnionymi w badaniu zachodzą istotne związki, których sprawdzenia nie podjęto się w ramach etapu I i II. Wyniki wszystkich analiz korelacyjnych zostały zestawione w tabeli 3.5.

Etap I: DFT x APM

W celu zbadania związku pomiędzy wynikami DFT i APM obliczono współczynnik korelacji r -Pearsona. Korelacja okazała się być istotna statystycznie ($r = 0,295$; $p = 0,044$). Wielkość współczynnika r pozwala wnosić, że mamy w tym wypadku do czynienia z niskim, dodatnim związkiem pomiędzy zmiennymi.

Etap II: DFT x NFCS, DFT x SEUA

Aby sprawdzić, czy wyniki DFT są istotnie powiązane z wynikiem ogólnym NFCS, ponownie obliczono współczynnik korelacji r -Pearsona. Związek pomiędzy tymi zmiennymi okazał

się być nieistotny statystycznie ($r = -0,179$; $p = 0,229$).

W celu zbadania związków pomiędzy wynikami DFT a wynikami uzyskanymi w podskalach NFCS posłużono się zarówno współczynnikiem r -Pearsona jak i r -Spearmana (wybór uzależniony był od kształtu rozkładu wyników uzyskanych w kontekście danej podskali testu NFCS). Wyniki testu DFT okazały się nie być istotnie powiązane z żadną z podskal NFCS (por. tabela 3.5).

Związki pomiędzy wynikami DFT a wynikami podskal SEUA zbadano obliczając wskaźniki r -Spearmana dla każdej z par wyników. Okazało się, że wyniki DFT są istotnie skorelowane z podskalą SEUA „prawda w kontekście świata społecznego”; w wymienionym przypadku uzyskano istotny umiarkowany, dodatni związek pomiędzy zmiennymi ($r_s = 0,332$; $p = 0,029$). Pozostałe związki okazały się być nieistotne statystycznie (por. tabela 3.5).

Etap III: pozostałe

Dodatkowe analizy zostały przeprowadzone z zastosowaniem zarówno testu r -Pearsona jak i r -Spearmana. W większości przypadków nie zaobserwowano statystycznie istotnych korelacji (por. tabela 3.5); wyjątkami były istotne związki pomiędzy podskalą SEUA „sądy związane z gustem” a dwiema podskalami NFCS: „zdecydowanie” (związek umiarkowany, dodatni: $r_s = 0,324$; $p = 0,028$) oraz „nietolerancja wieloznaczności” (związek umiarkowany, ujemny: $r_s = -0,308$; $p = 0,036$).

Tabela 3.5: Wyniki analiz korelacyjnych

test	1	2	3	4	5	6	7	8	9	10	11	12	13
1	-	$r = 0,295^*$	$r = -0,024$	$r_s = -0,216$	$r = 0,141$	$r_s = -0,122$	$r_s = 0,078$	$r = -0,092$	$r_s = -0,062$	$r_s = -0,014$	$r_s = -0,025$	$r_s = 0,186$	$r_s = -0,064$
2	-	-	$r = -0,179$	$r_s = -0,208$	$r = -0,107$	$r_s = -0,072$	$r_s = 0,073$	$r = -0,040$	$r_s = 0,007$	$r_s = 0,096$	$r_s = 0,001$	$r_s = 0,322^*$	$r_s = 0,115$
3	-	-	-	$r_s = 0,508^{**}$	$r = 0,687^{**}$	$r_s = 0,696^{**}$	$r_s = 0,204$	$r = 0,151$	$r_s = -0,228$	$r_s = -0,254$	$r_s = -0,287$	$r_s = 0,047$	$r_s = 0,016$
4	-	-	-	-	$r_s = 0,374^{**}$	$r_s = -0,358^*$	$r_s = -0,087$	$r_s = -0,234$	$r_s = -0,131$	$r_s = -0,061$	$r_s = -0,107$	$r_s = 0,055$	$r_s = 0,064$
5	-	-	-	-	-	$r_s = 0,422^{**}$	$r_s = -0,187$	$r = -0,195$	$r_s = -0,133$	$r_s = -0,231$	$r_s = -0,272$	$r_s = 0,163$	$r_s = 0,058$
6	-	-	-	-	-	-	$r_s = 0,037$	$r_s = -0,249$	$r_s = -0,308^*$	$r_s = -0,165$	$r_s = -0,057$	$r_s = 0,079$	$r_s = -0,108$
7	-	-	-	-	-	-	-	$r_s = -0,132$	$r_s = -0,052$	$r_s = 0,108$	$r_s = -0,004$	$r_s = -0,078$	$r_s = -0,002$
8	-	-	-	-	-	-	-	-	$r_s = 0,324^*$	$r_s = -0,200$	$r_s = -0,108$	$r_s = -0,023$	$r_s = 0,089$
9	-	-	-	-	-	-	-	-	-	$r_s = 0,139$	$r_s = 0,328^*$	$r_s = 0,058$	$r_s = -0,044$
10	-	-	-	-	-	-	-	-	-	-	$r_s = 0,204$	$r_s = -0,067$	$r_s = -1,116$
11	-	-	-	-	-	-	-	-	-	-	-	$r_s = -0,057$	$r_s = -0,091$
12	-	-	-	-	-	-	-	-	-	-	-	-	$r_s = 0,429^{**}$

* – korelacja istotna na poziomie 0,05

** – korelacja istotna na poziomie 0,01

Oznaczenia:

1 – APM;

2 – TGD;

3 – NFCS: ogólny;

4 – NFCS: preferowanie przewidywalności;

5 – NFCS: preferowanie porządku;

6 – NFCS: nietolerancja wieloznaczności;

7 – NFCS: zamkniętość umysłowa;

8 – NFCS: zdecydowanie;

9 – SEUA: sądy związane z gustem;

10 – SEUA: sądy związane z estetyką;

11 – SEUA: wartości moralne; 12 – SEUA: prawda w kontekście świata społecznego;

13 – SEUA: prawda w kontekście świata fizycznego.

4 Dyskusja wyników

4.1 Analiza rzetelności

Wyniki analiz rzetelności zaprezentowane w podrozdziale 3.3 upoważniają do stwierdzenia, że dla zdecydowanej większości narzędzi użytych w badaniu uzyskano co najmniej satysfakcjonujące poziomy wskaźników α Cronbacha bądź λ_2 Guttmana. Co również godne odnotowania, otrzymane wskaźniki są w dużej mierze spójne z tymi uzyskanymi w trakcie wcześniejszych badań z ich użyciem przeprowadzanych na polskich próbach (por. np. [Kossowska, Hanusz i Trejtowicz, 2012; Urbański, Paluszkiewicz i Urbańska, 2014; Żyłuk, 2016]). Szczególnie wysokie współczynniki odnotowano w przypadku APM oraz DFT (λ_2 Guttmana wynosząca, odpowiednio, 0,789 oraz 0,820). Istotnym wyjątkiem w zakresie obu powyższych kryteriów (satysfakcjonująca wielkość wskaźników oraz spójność wyników z raportowanymi wcześniej) są jednak wyniki analiz rzetelności testu SEUA. W kontekście testu SEUA – zastosowanego w opisywanym badaniu w wersji papier-ołówek – uzyskano niższe poziomy rzetelności aniżeli wtedy, gdy użyto SEUA w formie wywiadu [Żyłuk i in., 2016]. Jak pozwala wnosić treść tabeli 3.4, jedna z podskal testu SEUA (sądy związane z gustem) uzyskała nieakceptowalnie niski poziom spójności wewnętrznej.

4.2 Analiza korelacyjna

Analizy ukierunkowane na zbadanie powiązań pomiędzy płynnością w zakresie trudnej dedukcji a inteligencją płynną pozwalają na wniosek, że zmienne te są ze sobą pozytywnie skorelowane. Odnotowana korelacja, choć statystycznie istotna, jest jednak korelacją niską (współczynnik r nieco poniżej 0,3). Zgodnie z wynikami wcześniejszych badań nad związkami pomiędzy inteligencją płynną a dwoma typami dedukcji: trudną i łatwą [Urbański, Paluszkiewicz i Urbańska, 2014], uzyskany rezultat wskazywać może, że model inteligencji Catella-Horna-Carolla nie jest w stanie objąć niektórych aspektów umiejętności rozumowania dedukcyjnego. Wyczerpująca interpretacja uzyskanego rezultatu byłaby jednak możliwa jedynie pod warunkiem zebrania większej ilości danych i z uwzględnieniem dodatkowych parametrów charakteryzujących osoby badane.

W trakcie przyszłych badań nad związkami inteligencja płynna–trudna dedukcja konieczne byłoby wzięcie pod uwagę wpływu doświadczenia w zakresie nauki logiki – podobnie, jak uczyniono to w badaniu opisanym w raporcie Urbańskiego, Paluszkiewicz i Urbańskiej [2014]. W trakcie owego badania zaobserwowano, że biegłość w zakresie trudnych rozumowań dedukcyjnych (mierzonych *Testem Rozumowań Erotetycznych*), choć związana z poziomem inteligencji płynnej, zależy również od doświadczenia; co istotne, zależności o podobnym charakterze nie zaobserwowano w kontekście zadań angażujących zdolności w zakresie prostych rozumowań dedukcyjnych (*Test Rozumowań Dedukcyjnych*). Autorzy przywoływanego badania zaobserwowali, że w grupie wyćwiczonej w logice formalnej (kilka semestrów przedmiotów logicznych na studiach) nie było związku pomiędzy trudną dedukcją a inteligencją płynną, podczas gdy w grupach, które zrealizowały jedynie podstawowy program logiki na studiach, stwierdzono zachodzenie średnich i wysokich korelacji pomiędzy tymi zmiennymi.

Znamienne, że w grupie osób, które miały większe doświadczenie w zakresie nauki logiki uzyskano istotnie wyższe wyniki w teście trudnej dedukcji niż w grupach z mniejszym doświadczeniem tego rodzaju (zbadano dwie „niedoświadczone” grupy; jedna z grup nie różniła się pod względem poziom inteligencji płynnej od grupy z dużym doświadczeniem, druga z grup osiągała istotnie niższe wskaźniki inteligencji płynnej niż obie grupy; pomiar inteligencji płynnej dokonany był z zastosowaniem APM). Co także istotne, grupa wytrenowana w zakresie logiki oraz grupa niewytrenowana, ale na tym samym poziomie inteligencji, nie różniły się istotnie w zakresie wykonania *Testu Rozumowań Dedukcyjnych* (prosta dedukcja), przy czym obie te grupy uzyskały wyższe rezultaty w tym teście niż grupa osób niewytrenowanych w logice, ale o niższych wynikach w APM. Autorzy badania skonkludowali, że w kontekście grupy doświadczonej w logice ich wyższy poziom wykonania w zakresie trudnych problemów dedukcyjnych może efektem zadziałania czynnika nieobecnego w przypadku dwóch pozostałych grup. Wedle badaczy, wyniki przez nich uzyskane sugerują – jak zasygnalizowano wcześniej – że choć poziom wykonania zadań wchodzących w skład *Testu Rozumowań Erotetycznych* jest w pewnym zakresie związany z czynnikiem płynnym inteligencji ogólnej, ma on również związek z doświadczeniem uczestników – konkretnie, doświadczeniem w logiki formalnej jako przedmiotu nauczanego na studiach. Być może zatem zasadne jest wyróżnienie takich typów rozumowań dedukcyjnych, w przypadku których osiągnięcie pewnego, wysokiego pułapu wyników w testach je mierzących, wiązałoby się raczej z poziomem inteligencji skryształizowanej, aniżeli płynnej.

W badaniu opisywanym w niniejszym raporcie uczestniczyły osoby o zróżnicowanych poziomach w zakresie logicznej edukacji. Z uwagi na niewielką liczebność próby nie było możliwe wydzielenie podgrup, w obrębie których można byłoby przeprowadzić analizy korelacyjne wedle schematu zaproponowanego w przytaczanym powyżej raporcie. Przeprowadzenie badania w sposób analogiczny do powyżej opisanego umożliwiłoby stwierdzenie, czy, wielkość związku pomiędzy wynikami APM oraz wynikami DFT może różnić się w zależności od poziomu doświadczenia w zakresie logiki formalnej.

Inne wyjaśnienie – niekoniecznie alternatywne wobec powyższego – niskiego związku pomiędzy wynikiem DFT a APM, może być związane z przedmiotem pomiaru testu DFT. Niewykluczone, że zdolności dedukcyjne są tylko jedną z kilku charakterystyk poznawczych, o których poziomie można wnosić na podstawie wyniku *Testu Giętkości Dedukcyjnej*. Być może „zanieczyszczenie” pomiaru dokonywanego przy użyciu DFT spowodowało, że korelacja była niska. W przyszłości należałoby przeprowadzić badania ukierunkowane na zbadanie związków pomiędzy wynikami w DFT a innymi miarami poznawczymi takimi jak np. pamięć robocza, w celu dokładniejszego dookreślenia umiejętności, o których poziomie pozwala wnosić wynik DFT. Warto byłoby również zbadać powiązania pomiędzy wynikami DFT a innymi testami zdolności dedukcyjnych (np. prosta dedukcja – *Test Rozumowań Dedukcyjnych* czy trudna dedukcja – test *Style Rozumowania Dedukcyjnego* i *Test Rozumowań Erotetycznych*) [Paluszkiewicz, 2014; Urbański, Paluszkiewicz i Urbańska, 2014]. Pewnych informacji na temat przedmiotu pomiaru DFT dostarczyłoby również badanie z użyciem owego narzędzia przeprowadzone w formie wywiadu z uzupełniającym test uczestnikiem.

W przypadku analiz dotyczących powiązań DFT i NFCS (wyniku ogólnego i wyników pod-

skal) nie zaobserwowano zachodzenia jakichkolwiek istotnych korelacji. Próbuąc dociekać przyczyn takiego rezultatu, można odwołać się do obserwacji Kruganskiego i Webstera [1996], autorzy ci stwierdzili bowiem, że choć osoby cechujące się wysokim poziomem potrzeby domknięcia poznawczego często ograniczają swoje procesy poszukiwania i przetwarzania informacji, mierząc się z jakimś zadaniem, w pewnych szczególnych przypadkach, gdy wstępne przetwarzanie informacji nie pozwoli im na osiągnięcie domknięcia, ich wysoka potrzeba osiągnięcia tego poznawczego stanu może – paradoksalnie – wspierać bardziej rozległe i pogłębione analizy dostępnych informacji, ukierunkowane na wykształcenie stabilnej wiedzy. Niewykluczone, że brak związku pomiędzy wynikami DFT i NFCS był rezultatem zainicjowania powyżej opisanego mechanizmu w przypadku osób cechujących się wyższym poziomem potrzeby domknięcia.

Odnosząc się do wyników uzyskanych w trakcie trzeciej, dodatkowej, fazy analiz, należy zaznaczyć, że wynik wskazujący na brak związku pomiędzy poziomem inteligencji płynnej a poziomem potrzeby poznawczego domknięcia nie zaskakuje – jest on bowiem spójny z rezultatami wcześniej raportowanymi w literaturze [Kossowska, 2003; Kruglanski i Webster, 1996]. Wyjaśnienie owego braku powiązań nawiązuje zwykle do przytoczonych w poprzednim akapicie obserwacji Kruganskiego i Webstera [1996].

Nieco szerszego komentarza wymaga jednak seria wyników zebranych z użyciem testu SEUA. Wyniki analizy rzetelności oraz obserwacje procesu rozwiązywania testu przez uczestników badania pozwalają na wniosek, że pomimo dołożenia wszelkich starań ze strony osób opracowujących papierową wersję SEUA, metodą wywiadu jest wciąż bardziej odpowiednią formą przeprowadzania badania z użyciem owego narzędzia, aniżeli metoda papier-olówek.

Jak wspomniano wcześniej, w trakcie rozwiązywania testu SEUA, w przypadku każdej pozycji testowej uczestnicy badania mierzą się z zadaniem polegającym na ocenie, czy dwa zaprezentowane poglądy mogą być jednocześnie słuszne, i – jeżeli tak – czy któryś z nich może być bardziej słuszny. Odpowiedź na te pytania wymaga chwilowego zawieszenia własnych poglądów na temat tego, czego dotyczą owe poglądy – badany nie ma wybierać, który pogląd jest mu bliższy, pytany jest jedynie o swoje przekonania na temat możliwych relacji między tymi poglądami. Jest to zadanie szczególnie kłopotliwe w kontekście pytań o sądy dotyczące bardzo konkretnych sytuacji czy obiektów. W ich przypadku wpływ indywidualnych przekonań na tematy poruszane w prezentowanych zdaniach może być trudny do zahamowania. Nieco mniej problematyczne w tym kontekście były pozycje, na które składały się sądy o bardziej abstrakcyjnej treści – nie zawierające odniesień do szczególnych przedmiotów czy sytuacji, a określające jedynie, że dwa poglądy na dany temat są odmienne (np. para: „Robert sądzi, że pierwszy utwór, którego słuchali, jest lepszy.” oraz „Michał sądzi, że drugi utwór, którego słuchali, jest lepszy.” – nie jest w tym przypadku dokładnie określone, o jaki utwór chodzi). Szczegółowe omówienie tej problematyki znajduje się w artykule Żyłuk i współpracowników [2016].

W trakcie wywiadów przeprowadzanych z zastosowaniem SEUA wielu uczestników potrzebowało wsparcia ze strony przeprowadzającego wywiad w celu właściwego zrozumienia

instrukcji testu oraz pytań, jakie pojawiały się przy każdej z pozycji testowych. W przeważającej części sytuacji pomoc badacza polegała na przypominaniu, aby nie kierować się własnymi poglądami na poruszane tematy, ale aby, zgodnie z treścią pytań, oszacować możliwość jednoczesnej słuszności obu poglądów. W kontekście badania opisywanego w niniejszym raporcie, reagowanie w przypadku pojawienia się jakichkolwiek wątpliwości czy niejasności ze strony badanego nie było możliwe – głównie z uwagi na fakt, że uczestnicy badania nie zdawali sobie sprawy z tego, że w niewłaściwy sposób zinterpretowali zadanie, jakie przed nimi stoi, w związku z czym nie zgłaszali jakichkolwiek problemów badaczowi. Niemniej, na podstawie komentarzy jakie zostawiano po badaniu, uprawniony jest wniosek, że brak stałego kontaktu z badaczem mogącym korygować pomyłki badanego, skutkowało tym, że zadanie, jakie stało przed rozwiązującymi SEUA nie było, w istocie, właściwie zrozumiane.

Komentarze pozostawiane przez badanych; fakt, że jeden z uczestników nie uzupełnił testu we właściwy sposób oraz niesatysfakcjonujące poziomy rzetelności podskal SEUA (niższe niż uzyskane w kontekście badania metodą wywiadu) wpłynęły na decyzję o tym, aby nie interpretować wyników analizy korelacyjnej z zastosowaniem SEUA w wersji papier-olówek. W świetle powyższego, uzyskane wyniki, stwarzając dodatkowo duże trudności interpretacyjne, wydają się być artefaktami spowodowanymi złym zrozumieniem instrukcji testowych. Obecnie prowadzone prace skoncentrowane na udoskonaleniu testu SEUA w wersji papier-olówek. Celem jest opracowanie narzędzia, którego pozycje testowe oraz instrukcja będą na tyle zrozumiałe, aby w trakcie jego rozwiązywania nie było potrzebne ciągle wsparcie badacza. Poprawiona wersja narzędzia ma być złożona wyłącznie z pozycji testowych, które nie będą zawierać odniesień do konkretnych obiektów czy sytuacji.

SPIS ZAŁĄCZNIKÓW

Załącznik 1: Dane

Załącznik 2: Materiały wykorzystane w badaniu

- 2.1 Pisemna informacja o badaniu (APM, DFT, NFCS, SEUA)
- 2.2 Standaryzacja elementów sytuacji badania (APM, DFT, NFCS, SEUA) – Scenariusz przeprowadzenia badania wraz z instrukcją dla badającego

Załącznik 3: Materiały dodatkowe

- 3.1 Pisemna informacja o badaniu (DFT samodzielnie)
- 3.2 Standaryzacja elementów sytuacji badania (DFT samodzielnie) – scenariusz przeprowadzenia badania wraz z instrukcją dla badającego

BIBLIOGRAFIA

- Anastasi, A., Urbina, S. [1999]. *Testy psychologiczne*. A. Matczak (red.). Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Blair, C., [2010]. Fluid Cognitive Abilities and General Intelligence: A Life-Span Neuroscience Perspective. w: R. M. Lerner i W. F. Overton, W. F. (red.), *The Handbook of Life-Span Development. Cognition, Biology, and Methods*, John Wiley & Sons.
- Flanagan, D. [2008]. The Cattell-Horn-Carroll (CHC) theory of cognitive abilities. W: C. Reynolds, K. Vannest i E. Fletcher-Janzen (red.), *Encyclopedia of Special Education*, s. 368–386, John Wiley and Sons.
- Horn, J. L., Noll, J. [1997]. Human cognitive capabilities: Gf-Gc theory. W: D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (eds.). *Contemporary intellectual assessment: Theories, tests, and issues*, NY: Guilford, s. 53–91.
- Hornowska, E. [2007]. *Testy psychologiczne: teoria i praktyka..* Warszawa: Wydawnictwo naukowe „Scholar”.
- Guttman, L. [1945]. A basis for analysing test-retest reliability. *Psychometrika*, 10, s. 255–282.
- Hornowska, E. [2007]. *Testy psychologiczne: teoria i praktyka..* Warszawa: Wydawnictwo naukowe „Scholar”.
- Kossowska, M., Hanusz, K., Trejtowicz, M. [2012]. Skrócona wersja Skali Potrzeby Poznawczego Domknięcia. Dobór pozycji i walidacja skali. *Psychologia Społeczna*, 7(1), s. 89–90.
- Kossowska, M. [2003]. Różnice indywidualne w potrzebie poznawczego domknięcia. *Przegląd Psychologiczny*, 46(4), s. 355-374.
- Kruglanski, A., Webster, D. M. [1996]. Motivated closing of the mind: “Seizing” and “freezing”. *Psychological Review*, 103, s. 263–283.
- Kuhn, D., Cheney, R., Weinstock, M. [2000]. The development of epistemological understanding. *Cognitive development*, 15(3), s. 309–328.
- Paluszkiwicz, K. [2014]. *Polisylogisms – report*. Raport badawczy, Instytut Psychologii, Uniwersytet im. Adama Mickiewicza w Poznaniu.
- Peter, J., Rosman, T., Mayer, A. K., Leichner, N., Krampen, G. [2015]. Assessing epistemic sophistication by considering domain-specific absolute and multiplicistic beliefs separately. *British Journal of Educational Psychology*. 86, s. 204–221.
- Raven, J., Raven, J.C, Court, J.H. [2011]. *Podręcznik do Testu Matrycy Ravena oraz Skali Słownikowych (4: Wersja dla Zaawansowanych)*, Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego, Warszawa.

- Schneider, W., McGrew. [2012]. K. The Cattell-Horn-Carroll model of intelligence. In D. Flanagan, P. Harrison (eds.), *Contemporary intellectual assessment: Theories, tests, and issues* pp. 99–144, Guilford Press.
- Scott, W. A [1962]. Cognitive complexity and cognitive flexibility. *Sociometry*, 4, s. 405–414.
- Stenning, K., van Lambalgen, M. [2008]. *Human Reasoning and Cognitive Science*. Cambridge, MA: The MIT Press.
- Urbański, M., Paluszkiwicz, K., Urbańska. [2014]. *Deductive Reasoning and Learning: a Cross-Curricular Study*. Raport badawczy, Instytut Psychologii, Uniwersytet im. Adama Mickiewicza w Poznaniu.
- Żyluk., N. [2016]. *Test Giętkości Dedukcyjnej – raport z konstrukcji narzędzia*. Raport badawczy, Instytut Psychologii, Uniwersytet im. Adama Mickiewicza w Poznaniu.
- Żyluk, N., Karpe, K., Michta, M., Potok, W., Paluszkiwicz, K., Urbański, M. [2016]. Assessing Levels of Epistemological Understanding: The Standardized Epistemological Understanding Assessment (SEUA). *Topoi*, 2016.