

Polska adaptacja zadania Stroopa w warunku emocjonalnym do oceny preferencji pedofilnych – badania pilotażowe

Polish adaptation of emotional Stroop test in assessment of pedophilia – a pilot study

Wojciech Oronowicz-Jaśkowiak¹, Michał Lew-Starowicz^{1,2}

¹ Instytut Psychiatrii i Neurologii, III Klinika Psychiatryczna

² Centrum Medyczne Kształcenia Podyplomowego, Klinika Psychiatrii

Summary

Aim. A pilot study was conducted in order to construct a Polish adaptation of emotional Stroop test in assessment of pedophilia.

Material and method. The study consisted of two stages. The first stage involved creating test material by ranking words in adequate lists by competent experts. The second stage consisted of empirical verification of the principle of emotional Stroop test in a non-clinical population.

Results. Based on the assessment of five competent experts, words were ordered from the most to the least sexually arousing (Kendall's W from 0.368 to 0.693). Six ranked lists were obtained, and the competent experts were subsequently asked to assess whether these lists were suitable for the study (Lawshe's Content Validity Ratio from 0.6 to 1.0). Two categories of words were merged. Five ranked lists were obtained, and the competent experts were subsequently asked again to assess whether these lists were suitable for the study (Lawshe's Content Validity Ratio 1.0). The created lists of words were approved by all competent experts.

Based on the experimental study conducted on a non-clinical population, it was shown that, in accordance with the principle of the test, the mean response time for sexually related words was longer than for neutral words. The mean response time for children-related words did not differ significantly from response time for neutral words.

Conclusions. Based on the study with competent experts and conducted experiments, an initial Polish adaptation of the emotional Stroop test for diagnosis of pedophilia has been created. Further studies with persons with pedophilia are needed to implement the test in clinical setting.

Słowa kluczowe: pedofilia, kontrola uwagowa, zadanie Stroopa

Key words: pedophilia, attention control, Stroop task

Wprowadzenie

Jednym z zadań stosowanych w psychologii poznawczej jest zadanie Stroopa. W wersji oryginalnej zadaniem uczestnika badania jest jak najszybsze nazywanie kolorów czcionek wyrazów wyświetlanych na ekranie komputera. W warunku zgodnym wyświetlane jest słowo odpowiadające kolorowi czcionki (np. napis „niebieski” wyświetlony niebieską czcionką), natomiast w warunku niezgodnym kolor czcionki nie odpowiada znaczeniu danego słowa (np. napis „czerwony” wyświetlony czcionką zieloną). Wielokrotnie wykazywano, że w wypadku niezgodności koloru czcionki oraz znaczenia poszczególnych słów czas reakcji jest znacząco dłuższy [1].

W późniejszym czasie prezentowano wiele wersji zmodyfikowanego zadania Stroopa, które poza słowami odnoszącymi się do nazw kolorów zawierały słowa potencjalnie nacechowane emocjonalnie dla poszczególnych grup osób – alkoholu [2] czy też depresji [3]. Uzyskiwano wyniki świadczące o tym, że osoby, które były klasyfikowane przez autorów badania jako potencjalnie przejawiające stosunek emocjonalny do słów odnoszących się do specyficznych przedmiotów lub sytuacji, potrzebowały istotnie statystycznie więcej czasu na podjęcie decyzji co do nazwania koloru odpowiedniego słowa.

Co ważne z punktu widzenia niniejszej pracy, prezentowano również zmodyfikowaną wersję zadania Stroopa w badaniach sprawców przestępstw seksualnych. Ciardha i Gormley [4] przeprowadzili badanie na grupie 24 sprawców przestępstw seksualnych przeciwko małoletnim, uzyskując wyniki świadczące o różnicy w czasie przetwarzania bodźców odnoszących się do dzieci w porównaniu z grupą kontrolną. Price i wsp. [5] dowiedli, że grupa sprawców przestępstw seksualnych prezentuje dłuższy czas reakcji w wypadku przetwarzania słów nacechowanych seksualnie. Podobne badania przeprowadzili Smith i Waterman [6], Price i Hanson [7] oraz van Leeuwen i wsp. [8]. Interpretacja uzyskiwanych wyników jest jednak utrudniona ze względu na niewystarczającą liczbę publikacji naukowych, niejednorodność grup badawczych oraz nieuwzględnienie w badaniach porównań między osobami ze zdiagnozowanymi zaburzeniami preferencji seksualnych a osobami bez tego zaburzenia. Wyników prezentowanych badań nie można ponadto wykorzystywać w praktyce sądowo-seksuologicznej, ponieważ nie przeprowadzono normalizacji odpowiadającej standardom stawianym narzędziom psychologicznym. Warto jednak dodać, że według Price i Hansona [7] badania prowadzące do stworzenia takiego narzędzia oraz jego zastosowania w praktyce wydają się obiecujące. Jako że potencjalnym zastosowaniem zmodyfikowanego zadania Stroopa mógłby być obszar sądowo-seksuologiczny, wnioski zawarte w diagnozie powinny się opierać na szczególnie restrykcyjnych badaniach. Niezbędne są dalsze badania walidacyjne zmierzające do określenia jego właściwości psychometrycznych.

Uzasadnione wydaje się przeprowadzenie badań pilotażowych zmierzających do utworzenia polskich list słów wchodzących w skład zadania Stroopa w warunku emocjonalnym. W oryginalnym badaniu Smitha i Watermana [6] stworzono listy słów, opierając się na rekomendacjach psychologów sądowych oraz pracowników służby więziennej pracujących w oddziałach terapeutycznych dla sprawców przestępstw

seksualnych. Brak jest szczegółowych informacji o procesie konstruowania listy słów, co z metodologicznego punktu widzenia nie powinno być dopuszczalne.

W badaniu Price i Hansona [7] utworzono listy słów na podstawie ocen dokonywanych przez pięciu sędziów kompetentnych, którzy decydowali, czy poszczególne słowa mogą zostać zaklasyfikowane do badania. W skład grupy sędziów kompetentnych wchodziłi wszakże nie eksperci z zakresu seksuologii, tylko studenci, ponadto do tego typu badań właściwe jest wykorzystanie współczynnika trafności treściowej Lawshe'a (CVR), jednak autorzy nie odwołują się w swoich badaniach do tej statystyki. Dodatkowo zaklasyfikowano do dalszych badań słowa, które zgodnie z wartością krytyczną dla pięciu sędziów kompetentnych [9] nie powinny być zostać zaklasyfikowane, ponieważ wartość CVR przy niezgodności dwóch z pięciu sędziów kompetentnych jest niższa niż próg $CVR = 1,00$. Ponadto Price i Hanson [7] nie szeregowali wstępnie wygenerowanej listy słów pod względem siły wywoływanego podniecenia u osób z preferencjami pedofilnymi. Wskazywano jedynie, czy dane słowo powinno zostać zaklasyfikowane. W badaniach Smitha i Watermana [6] brak jest informacji o stosowanej metodzie doboru słów (szeregowania lub wyboru kategorialnego). Konieczne wydaje się zatem rozpoczęcie procesu tworzenia nowych list, jako że dotychczasowe badania prezentowały w tym zakresie istotne błędy metodologiczne. Narzędzie do użytku klinicznego powinno spełniać wysokie kryteria metodologiczne, w tym na etapie procesu konstrukcji narzędzia.

Uzasadnione ponadto jest zweryfikowanie, czy polska adaptacja zadania Stroopa w warunku emocjonalnym odpowiada charakterystyce uzyskanych wyników w badaniach anglojęzycznych, w szczególności w wypadku stosowania słów wulgarnych oraz niewulgarnych. Ograniczeniem w interpretacji wyników dotychczasowych badań jest brak podziału między słowami nacechowanymi seksualnie niewulgarnymi oraz wulgarnymi słowami nacechowanymi seksualnie. Wnioskuje się o dłuższych czasach reakcji w wypadku słów nacechowanych seksualnie, przy czym trudno określić, czy przyczyną dłuższego czasu przetwarzania może być niespodziewane wystąpienie słowa wulgarnego czy też większe znaczenie słów nacechowanych seksualnie dla uczestnika badania. Z tego powodu istotne wydaje się wprowadzenie rozróżnienia między tymi grupami słów.

Hipotezy

Etap I konstrukcji narzędzia

Ze względu na charakter eksploracyjny badania z udziałem sędziów kompetentnych nie postawiono hipotez badawczych. Celem badań było uzyskanie listy słów uszeregowanych przez sędziów kompetentnych od najmniej podniecających do najbardziej podniecających dla osób ujawniających preferencje pedofilne.

Etap II konstrukcji narzędzia

Jeżeli poszczególne listy słów zostały poprawnie zaadaptowane do warunków polskich, to powinno się oczekiwać podobnych wyników. Jeżeli zatem stwierdzono, że osoby skazane za przestępstwa seksualne przeciwko dzieciom prezentują dłuższy czas reakcji w wypadku słów odnoszących się do dzieci [4, 7], to należałoby oczekiwać, że osoby z populacji nieklinicznej nie będą prezentowały podobnego wzorca wykonania zadania. Analogicznie, jeżeli stwierdzono, że osoby samoidentyfikujące się jako osoby z preferencjami pedofilnymi prezentują dłuższy czas reakcji w wypadku słów odnoszących się do dzieci [4], to również należałoby oczekiwać, że osoby z populacji nieklinicznej nie będą prezentowały podobnego wzorca wykonania zadania. W przeciwnym razie mogłaby zostać zakwestionowana trafność zadania Stroopa do oceny preferencji pedofilnych.

Na podstawie przedstawionego przeglądu literatury postawione zostały hipotezy badawcze odnoszące się do średnich skorygowanych czasów reakcji na słowa z poszczególnych kategorii zadania Stroopa w warunku emocjonalnym. Oczekiwano, że średni skorygowany czas reakcji na określenie koloru słów oznaczających nazwy kolorów (H1) oraz na wulgaryzmy (H2) będzie dłuższy w porównaniu z czasem reakcji na słowa związane z dziećmi.

Material i metoda

Etap I konstrukcji narzędzia

W celu określenia listy słów będących w opinii ekspertów najlepszymi bodźcami w polskiej adaptacji zadania Stroopa w warunku emocjonalnym przeprowadzono badanie z udziałem sędziów kompetentnych w dziedzinie seksuologii. Proces wyboru ostatecznej listy słów składał się z trzech etapów: (1) utworzenie wstępnej listy słów; (2) ocena sędziów kompetentnych przez szeregowanie słów; oraz (3) ponowna ocena sędziów kompetentnych przez akceptację lub odrzucenie stworzonych list słów.

Pierwszy etap badania pierwszego polegał na wygenerowaniu listy słów, które w dalszej kolejności zostały przekazane do oceny przez sędziów kompetentnych. Listy słów zostały stworzone na podstawie puli słów-bodźców przetłumaczonych na język polski przez Kasparka (2015), stosowanej wcześniej w badaniach Price i Hansona [7], oraz na podstawie materiału ze słownika wyrazów bliskoznacznych.

Drugi etap polegał na poleceniu uszeregowania poszczególnych słów. Zadaniem sędziów kompetentnych było uszeregowanie słów od najbardziej pobudzających dla osób z preferencjami pedofilnymi do najmniej pobudzających, dla każdej z grup słów osobno. Grupa uczestników badania składała się z pięciu sędziów kompetentnych, w tym: dwóch psychologów – seksuologów klinicznych certyfikowanych przez Polskie Towarzystwo Seksuologiczne (PTS), jednego lekarza specjalistę seksuologa i seksuologa klinicznego PTS oraz dwóch psychologów, absolwentów studiów podyplomowych z zakresu seksuologii klinicznej, będących w trakcie procesu certyfikacji PTS. Sędziowie byli zapraszani osobiście do wzięcia udziału w badaniu.

Trzeci etap polegał na ocenie, czy utworzone listy słów mogą zostać zaakceptowane do dalszych badań. Grupę uczestników badania tworzył ten sam zespół sędziów kompetentnych. Zadaniem sędziów kompetentnych było podjęcie kategorycznej decyzji dotyczącej tego, czy utworzone w drugim etapie listy słów będą właściwe do badania. Podobnie jak w etapie drugim sędziowie dokonywali ocen dla każdej z grup osobno.

Etap II konstrukcji narzędzia

W celu stwierdzenia różnic między skorygowanymi średnimi czasami reakcji przeprowadzono badanie empiryczne polegające na wykonaniu kolejno dwóch zadań różniących się zestawem kategorii słów, przy czym każde zadanie obejmowało łącznie taką samą liczbę słów (tab. 1).

Tabela 1. Listy słów, które były prezentowane uczestnikom badania w poszczególnych zadaniach

Zadanie pierwsze	Zadanie drugie
Lista słów „neutralne” (n = 25)	Lista słów „neutralne” (n = 25)
Lista słów „kolorowe” (n = 25)	Lista słów „kolorowe” (n = 25)
Lista słów „pedofilne – nowe” (n = 15)	Lista słów „pedofilne – nowe” (n = 15)
Lista słów „ogólnoseksualne – wulgarne” (n = 15)	Lista słów „ogólnoseksualne – łagodne” (n = 15)
Lista słów „pedofilne – chłopcy” (n = 15)	Lista słów „pedofilne – dziewczynki” (n = 15)

Grupę uczestników badania stanowiło 25 osób ($M = 37,25$ roku; $SD = 15,56$), w tym 16 kobiet ($M = 39,33$ roku; $SD = 14,42$) oraz 9 mężczyzn ($M = 33,78$; $SD = 16,72$). Zastosowano nielosową metodę *ad hoc* rekrutacji uczestników badania do próby. W grupie badanej 22 osoby (88%) miały wyższe wykształcenie, natomiast trzy osoby (12%) zadeklarowały wykształcenie średnie.

Badanie prowadzone było z użyciem komputera przenośnego. Instrukcja do badania oraz właściwe zadanie były wyświetlane na ekranie komputera. Zadaniem uczestnika badania było wciśnięcie klawisza odpowiadającego kolorowi czcionki wyświetlanego słowa. Instrukcja dla uczestnika badania była taka sama w obydwu zadaniach:

W czasie kolejnych zadań zobaczysz słowa wyświetlone w różnych kolorach. Twoim zadaniem jest wskazanie koloru, którym zostało napisane dane słowo, ignorując co dane słowo właściwie znaczy. Wskaż kolor danych słów przyciskając odpowiedni klawisz: d dla czerwonych słów; f dla zielonych słów; j dla niebieskich słów; k dla żółtych słów. Spróbuj odpowiedzieć tak szybko i dokładnie, jak tylko potrafisz, ponieważ będzie mierzony czas odpowiedzi. Jeżeli popełnisz błąd, czerwony X zostanie wyświetlony na ekranie. Kontynuuj, kiedy będziesz gotowy na wyświetlenie przykładu.

Przed rozpoczęciem właściwego zadania uczestnik badania był proszony o wykonanie sesji treningowej polegającej na naciśnięciu odpowiedniego klawisza w odpowiedzi na kolor czcionki dziesięciu liczebników. Uczestnik badania decydował o okre-

śleniu koloru danego słowa przez naciśnięcie odpowiedniego klawisza na klawiaturze komputerowej. Wyniki z tej części nie były rejestrowane. Celem przeprowadzenia sesji treningowej było zapoznanie uczestników badania z procedurą oraz nabycie wprawy.

Po zakończeniu sesji treningowej uczestnik badania przechodził automatycznie do właściwego zadania. Uczestnikom badania prezentowane były wszystkie słowa wchodzące w skład poszczególnych kategorii słów z poleceniem jak najszybszego określenia koloru danego słowa. Słowa wyświetlane były w losowej kolejności, również stosowane kolory czcionek były ustalane losowo. Lista słów „neutralnych” obejmowała słowa pozbawione znaczenia emocjonalnego, natomiast lista słów „kolorowych” obejmowała słowa oznaczające kolory (zob. Załącznik 1). Lista słów „ogólnoseksualne – wulgarne” (zob. tab. 3) oraz „ogólnoseksualne – łagodne” (zob. tab. 2) obejmowała słowa związane z ludzką seksualnością. Listy słów „pedofilne – nowe” (zob. tab. 10), „pedofilne – chłopcy” (zob. tab. 6) oraz „pedofilne – dziewczynki” (zob. tab. 5) obejmowały słowa związane z dziećmi oraz dzieciństwem. Każda z list obejmowała po 15 słów wyróżnionych na podstawie badań z udziałem sędziów kompetentnych.

Czas reakcji na poszczególne słowa i liczba błędów były zapisywane automatycznie przez program do prowadzenia badań z zakresu psychologii – Inquist.

Wyniki

Etap I konstrukcji narzędzia

Na podstawie przedstawionych źródeł wygenerowano listę 180 słów, przyporządkowanych następnie do sześciu kategorii, w tym każda kategoria słów zawierała 30 słów. Poszczególne słowa zostały zgrupowane wokół podobnych zakresów tematycznych – słów określających: stosunki seksualne (listy nr 1 i 2), dzieci oraz dzieciństwo (listy nr 3 i 6), dziewczynki (lista nr 4) oraz chłopców (lista nr 5).

W celu określenia zgodności sędziów kompetentnych dla każdej z grup osobno obliczono współczynnik zgodności W Kendalla. Uszeregowane słowa zostały przedstawione w tabelach 2–7. W tabeli 8 zaprezentowano współczynniki zgodności W Kendalla dla poszczególnych grup słów. W dalszej kolejności, w celu określenia zgodności sędziów kompetentnych dla każdej z grup osobno, obliczono współczynnik trafności treściowej Lawshe’a (CVR). Wyniki zostały przedstawione w tabeli 9.

Tabela 2. Lista słów wchodzących w skład kategorii „ogólnoseksualne – łagodne” uszeregowana od najniższej do najwyższej rangi

Lista nr 1					
Lp.	Słowo	Średnia ranga*	Lp.	Słowo	Średnia ranga
1	wilgotna	6	16	dotykać	16,2
2	pieścić	7,2	17	wagina	16,2
3	rozbierać	8	18	majtki	16,4
4	seks	8,8	19	striptiz	16,8
5	rozebrana	9	20	masować	17,8

dalszy ciąg tabeli na następnej stronie

6	kusić	9,2	21	wibrator	18
7	jęczeć	9,4	22	stringi	18
8	ssać	9,6	23	flirt	18,8
9	orgazm	10	24	kochanka	19,8
10	pośladki	10,4	25	masturbacja	22,8
11	podniecać	11,6	26	erekcja	22,8
12	biust	13,2	27	odbyt	23,4
13	dziewica	14,8	28	prostyutka	24,4
14	lizać	15	29	miłość	26,2
15	pocałunek	16	30	czochrać	29,2

* Im niższa ranga, tym dane słowo oceniane było przez sędziów kompetentnych jako bardziej podniecające, ponieważ sędziowie szeregowali wszystkie trzydzieści słów od wartości 1 do 30, gdzie 1 oznaczało słowo najbardziej podniecające, a 30 słowo najmniej podniecające.

Tabela 3. Lista słów wchodzących w skład kategorii „ogólnoseksualne – wulgarne” uszeregowana od najniższej do najwyższej rangi

Lista nr 2					
Lp.	Słowo	Średnia ranga	Lp.	Słowo	Średnia ranga
1	cipa	4	16	jebać	17,4
2	oral	4,2	17	ślina	17,8
3	anal	4,8	18	bzykać	18
4	mokra	5,6	19	brać	18,2
5	cycki	6,6	20	dziwka	18,2
6	dupa	6,8	21	twardy	18,4
7	ruchać	7,6	22	ostry	19,4
8	posuwać	8	23	dupczyć	20,6
9	szpara	8,4	24	włosy	22,4
10	walić	12,2	25	dmuchać	22,4
11	spuścić	13,4	26	grzmocić	23,8
12	pieprzyć	14,2	27	zapinać	24
13	rznąć	14,8	28	rypać	26,6
14	kutas	15,2	29	rąbać	27
15	dymać	16,4	30	chędożyć	28,6

Tabela 4. Lista słów wchodzących w skład kategorii „pedofilne – łagodne” uszeregowana od najniższej do najwyższej rangi

Lista nr 3					
Lp.	Słowo	Średnia ranga	Lp.	Słowo	Średnia ranga
1	dziewczynka	4,6	16	buzia	16,2
2	chłopczyk	5,6	17	pieluszki	17,2
3	pupa	5,8	18	kotek	17,8
4	dzieci	7	19	dzieciństwo	18
5	zabawa	9	20	karuzela	19,2
6	tajemnica	9,2	21	huśtawka	19,6
7	myszka	9,6	22	przewijanie	19,6
8	zabawki	12,2	23	kaczuszka	19,8
9	niewinność	12,8	24	śliniaczek	20
10	sekret	12,9	25	kredki	21,2
11	kicia	13,4	26	śpioszki	21,6
12	przedszkole	13,8	27	zależność	21,6
13	podstawówka	14	28	lekcje	22,2
14	szkoła	15,4	29	opieka	24,8
15	smoczek	15,6	30	troskliwość	25,3

Tabela 5. Lista słów wchodzących w skład kategorii „pedofilne – dziewczynki” uszeregowana od najniższej do najwyższej rangi

Lista nr 4					
Lp.	Słowo	Średnia ranga	Lp.	Słowo	Średnia ranga
1	szparka	4,4	16	gówniara	15
2	cipcia	5	17	smarkuła	15,4
3	laleczka	7	18	małolata	16,8
4	dziurka	7,4	19	niedojrzała	17,2
5	słodka	9,8	20	sisiumajtką	17,2
6	dziewczynka	11	21	szczeniara	17,6
7	sisia	11,2	22	zależna	19,2
8	panieneczka	11,4	23	misia	19,9
9	panienka	13,6	24	nastolatka	19,9
10	warkoczyki	13,6	25	młoda	20,8
11	figlarka	13,8	26	sukienka	21,2
12	smarkatka	14,4	27	kilkulatka	21,8

dalszy ciąg tabeli na następnej stronie

13	niewinna	14,4	28	dzieweczka	23,2
14	córeczka	14,8	29	młodociana	24,4
15	księżniczka	14,8	30	śmietanka	28,8

Tabela 6. Lista słów wchodzących w skład kategorii „pedofilne – chłopcy” uszeregowana od najniższej do najwyższej rangi

Lista nr 5					
Lp.	Słowo	Średnia ranga	Lp.	Słowo	Średnia ranga
1	chłopczyk	5,6	16	młokos	16,4
2	chłopaczek	7	17	synek	16,8
3	smarkacz	7,2	18	żołnierzyki	17,2
4	siusiak	7,6	19	młodzieniec	17,4
5	małolat	8	20	młodociany	17,8
6	malec	10,4	21	niedojrzały	18,6
7	siusiumajtek	11,4	22	kilkulatek	19,2
8	zależny	11,6	23	młodzian	20
9	niewinny	12,2	24	wyrostek	20,4
10	smarkaty	12,4	25	młody	21
11	chłoptaś	12,6	26	nastolatek	21,2
12	gówniarz	14,4	27	ślodki	21,8
13	łobuz	14,8	28	dorastający	22,6
14	szczeniak	15,6	29	niepełnoletni	23,4
15	synuś	15,8	30	troskliwy	24,6

Tabela 7. Lista słów wchodzących w skład kategorii „pedofilne – ogólne” uszeregowana od najniższej do najwyższej rangi

Lista nr 6					
Lp.	Słowo	Średnia ranga	Lp.	Słowo	Średnia ranga
1	pupa	5,8	16	dzieciństwo	14,8
2	dzieci	6,2	17	dzidzia	15,8
3	przedszkolaczek	8,8	18	bobas	16,8
4	rączki	9,2	19	śpioszki	17,4
5	tajemnica	10	20	lulać	17,6
6	maluch	10,2	21	bawić	18
7	kolanka	10,4	22	szkoła	18
8	dzieciak	12,2	23	smoczek	19

dalszy ciąg tabeli na następnej stronie

9	zabawki	12,6	24	pieluszki	19
10	zabawa	13	25	mleczko	19,4
11	tulić	13	26	lekcje	22
12	przedszkole	13,4	27	opieka	22,8
13	dzidzius	14,2	28	kredki	24
14	tatus	14,2	29	śliniaczek	24,4
15	podstawówka	14,4	30	skórka	28,4

Tabela 8. Statystyka testu W Kendalla dla poszczególnych list słów

Numer listy słów	Statystyka testu
1	W Kendalla = 0,486; chi kwadrat = 70,517; df = 29; p < 0,001
2	W Kendalla = 0,693; chi kwadrat = 100,489; df = 29; p < 0,001
3	W Kendalla = 0,431; chi kwadrat = 62,565; df = 29; p < 0,001
4	W Kendalla = 0,424; chi kwadrat = 61,492; df = 29; p < 0,001
5	W Kendalla = 0,368; chi kwadrat = 53,320; df = 29; p = 0,004
6	W Kendalla = 0,385; chi kwadrat = 55,885; df = 29; p = 0,002

Tabela 9. Statystyka trafności treściowej Lawshe'a dla poszczególnych list słów

	Lista 1	Lista 2	Lista 3	Lista 4	Lista 5	Lista 6
Liczba udzielonych odpowiedzi	5	5	5	5	5	5
Liczba odpowiedzi pozytywnych	5	5	4	5	5	4
Liczba odpowiedzi negatywnych	0	0	1	0	0	1
CVR	1,00	1,00	0,60	1,00	1,00	0,60

Współczynnik trafności treściowej Lawshe'a wyniósł dla czterech grup słów 1,00, natomiast dla dwóch grup wyniósł 0,6, przekraczając wartość krytyczną pozwalającą na uznanie danej oceny za akceptowalną dla pięciu sędziów kompetentnych [9]. Czterej sędziowie kompetentni uznali, że utworzone listy słów mogą zostać wykorzystane w dalszych badaniach empirycznych. Jeden z sędziów uznał, że kategorie słów „pedofilne – ogólne” oraz „pedofilne – łagodne” nie powinny zostać uwzględnione w późniejszych badaniach, ponieważ obejmują podobny zakres tematyczny, ponadto występują w nich powtarzające się słowa. Zgodnie z sugestią sędziego kompetentnego połączono kategorie słów „pedofilne – łagodne” oraz „pedofilne – ogólne”, tworząc nową kategorię słów „pedofilne – nowe” (lista nr 6b). Lista słów z tej kategorii obejmowała słowa ocenione jako najbardziej podniecające z połączonych kategorii (zob. tab. 8).

Tabela 10. Połączone listy słów wchodzących w skład jednej listy uszeregowanej od najniższej do najwyższej rangi * 15 najbardziej podniecających słów, z wyłączeniem słów powtarzających się.

Lista nr 6b					
Lp.	Słowo	Średnia ranga	Lp.	Słowo	Średnia ranga
1	dziewczynka	5,8	31	smoczek	15,6
2	chłopczyk	6,2	32	dzidzia	15,8
3	pupa*	8,8	33	buzia	16,2
4	pupa	9,2	34	bobas	16,8
5	dzieci*	10	35	pieluszki	17,2
6	dzieci	10,2	36	śpioszki	17,4
7	przedszkolaczek*	10,4	37	lulać	17,6
8	zabawa*	12,2	38	kotek	17,8
9	tajemnica*	12,6	39	dzieciństwo	18
10	rażki*	13	40	bawić	18
11	myszka*	13	41	szkoła	18
12	tajemnica	13,4	42	smoczek	19
13	maluch*	14,2	43	pieluszki	19
14	kolanka*	14,2	44	karuzela	19,2
15	zabawki*	14,4	45	mleczek	19,4
16	dzieciak*	12,2	46	huśtawka	19,6
17	zabawki	12,6	47	przewijanie	19,6
18	niewinność*	12,8	48	kaczuszka	19,8
19	sekret*	12,9	49	śliniaczek	20
20	zabawa	13	50	kredki	21,2
21	tulić*	13	51	śpioszki	21,6
22	kicia*	13,4	52	zależność	21,6
23	przedszkole	13,4	53	lekcje	22
24	przedszkole	13,8	54	lekcje	22,2
25	podstawówka	14	55	opieka	22,8
26	dzidzius	14,2	56	kredki	24
27	tatus	14,2	57	śliniaczek	24,4
28	podstawówka	14,4	58	opieka	24,8
29	dzieciństwo	14,8	59	troskliwość	25,3
30	szkoła	15,4	60	skórka	28,4

Po utworzeniu nowej kategorii słów wszystkie listy przedstawiono do ponownej oceny sędziów kompetentnych. W celu określenia zgodności sędziów kompetentnych jeszcze raz dla każdej z grup osobno obliczono współczynnik trafności treściowej Lawshe'a. Wyniki zostały przedstawione w tabeli 11.

Tabela 11. Statystyka trafności treściowej Lawshe'a dla poszczególnych list słów

	Lista 1	Lista 2	Lista 3	Lista 4	Lista 5	Lista 6b
Liczba udzielonych odpowiedzi	5	5	5	5	5	5
Liczba odpowiedzi pozytywnych	5	5	5	5	5	5
Liczba odpowiedzi negatywnych	0	0	0	0	0	0
CVR	1,00	1,00	1,00	1,00	1,00	1,00

Etap II konstrukcji narzędzia

Wzorując się na badaniach Price i wsp. [5], przed przeprowadzeniem obliczeń skorygowano czas udzielania odpowiedzi przez odjęcie średnich czasów odpowiedzi w poszczególnych listach słów („kolorowe”, „pedofilne – nowe”, „ogólnoseksualne – wulgarne”, „pedofilne – chłopcy”, „ogólnoseksualne – łagodne”, „pedofilne – dziewczynki”) od średniego czasu odpowiedzi w liście słów neutralnych.

Zadanie 1

Przeprowadzono analizę wariancji ANOVA w celu określenia, czy istnieją różnice między skorygowanym średnim czasem odpowiedzi na dane kategorie słów w odpowiednich warunkach badawczych. Na podstawie testu Mauchly'ego stwierdzono, że nie spełniono założenia o sferyczności (W Mauchly'ego = 0,528; $p = 0,013$), wobec czego przeprowadzono obliczenia z poprawką Greenhouse'a-Geissera. Zaobserwowano różnice istotne statystycznie ($F = 4,249$; $p = 0,018$), w związku z czym przeprowadzono testy *post hoc* LSD. Wykazano, że średni skorygowany czas odpowiedzi na słowa z kategorii „wulgarne” był dłuższy od kategorii „pedofilne – chłopcy” (różnica średnich = 10740,1 ms; $p = 0,033$; 95% CI (956,67 ms, 20503,57 ms)). Średni skorygowany czas odpowiedzi na słowa z kategorii „kolorowe” był dłuższy od kategorii „pedofilne – nowe” (różnica średnich = 6823,40 ms; $p = 0,044$; 95% CI (185,40 ms, 13461,40 ms)) oraz dłuższy od kategorii „pedofilne – chłopcy” (różnica średnich = 11384,60 ms; $p = 0,006$; 95% CI (3580,24 ms, 19188,96 ms)).

Zadanie 2

Przeprowadzono analizę wariancji ANOVA w celu określenia, czy istnieją różnice między skorygowanym średnim czasem odpowiedzi na dane słowa w odpowiednich warunkach badawczych. W związku z brakiem różnic istotnych statystycznie ($p < 0,05$) nie przeprowadzono dalszych analiz statystycznych.

Dyskusja

Etap I konstrukcji narzędzia

Przeprowadzono badanie polegające na uszeregowaniu słów od najbardziej do najmniej podniecających przez sędziów kompetentnych z zakresu seksuologii. Celem badania było stworzenie odpowiednich kategorii słów do polskiej adaptacji zadania Stroopa do oceny preferencji pedofilnych. Początkowo utworzono listę 180 słów pogrupowanych w sześć kategorii.

Uzyskane listy słów mają posłużyć za materiał wyjściowy do stworzenia list słów możliwych do wykorzystania w dalszych pracach badawczych. Do ograniczeń badania należą jednak niskie współczynniki zgodności sędziów kompetentnych wobec uszeregowanych list słów na etapie drugim. Zasadniczo wszystkie współczynniki zgodności były niskie z wyjątkiem tego dla listy słów uwzględniającej słowa ogólnoseksualne. Z powodu niskiej czułości uszeregowanych słów niemożliwe jest zatem dokładne wnioskowanie o różnicach między rangami poszczególnych słów. W związku z tym w kolejnych pracach badawczych zaplanowano użycie jedynie połowy najwyżej ocenionych słów z poszczególnych pięciu list.

Do ograniczeń badania można zaliczyć fakt dokonywania oceny słów jako pobudzających dla osób z preferencjami pedofilnymi przez ekspertów z zakresu seksuologii, a nie przez osoby z preferencjami pedofilnymi. Wydaje się, że ocena słów przez osoby o preferencjach pedofilnych mogłaby być bardziej trafna, ponieważ nie opierałaby się na wiedzy ekspertów o preferencjach pedofilnych, tylko na rzeczywistych odczuciach osób z preferencjami pedofilnymi. Należy jednak zauważyć, że rekrutowanie do tego typu badania osób z preferencjami pedofilnymi mogłoby zostać uznane za nieetyczne, a dokonywana przez nie ocena mogłaby być zafałszowana, szczególnie w wypadku osób, które są w trakcie terapii z racji dokonania czynu w związku z zaburzeniami preferencji seksualnych.

Etap II konstrukcji narzędzia

W tym etapie badania określono różnice między średnimi skorygowanymi czasami odpowiedzi. Zgodnie z oczekiwaniami stwierdzono, że uczestnicy badania z populacji nieklinicznej dłużej podejmowali decyzję dotyczącą określenia koloru czcionki słów o zabarwieniu emocjonalnym, tj. słów oznaczających kolory (H1) oraz wulgaryzmy (H2), w porównaniu z czasem reakcji na słowa odnoszące się do dzieciństwa oraz dzieci.

Uzyskane wyniki korespondują z wcześniejszymi doniesieniami, zgodnie z którymi osoby z grupy kontrolnej prezentują krótszy czas reakcji w wypadku słów odnoszących się do dzieci, w porównaniu z osobami odbywającymi karę z powodu dokonania przestępstw seksualnych wobec dzieci, przy czym autorzy nie wprowadzili wystarczającej informacji o diagnozie zaburzeń preferencji seksualnych [4, 7].

Osoby włączone do badania z populacji nieklinicznej, z założenia bez preferencji pedofilnych, uzyskiwały w pierwszym zadaniu krótszy czas reakcji w wypadku słów odnoszących się do dzieciństwa w porównaniu z czasem reakcji na słowa o znaczeniu

emocjonalnym. Na szczególną uwagę zasługuje natomiast brak obserwacji analogicznego efektu w wypadku zadania drugiego, następującego od razu po wykonaniu zadania pierwszego. Uzyskane wyniki sugerują, że wykonanie zadania Stroopa w warunkach emocjonalnych po raz pierwszy znacząco wpływa na wyniki powtórnego wykonania zadania. Niemożliwe jest natomiast określenie, czy brak zaobserwowanego efektu jest wynikiem prezentacji listy słów „ogólnoseksualne – łagodne” w drugiej kolejności, czy też jest to efekt braku reakcji emocjonalnej uczestników badania na prezentowane słowa.

Z powyższych obserwacji wyciągnąć można co najmniej dwa wnioski. Po pierwsze, niewskazane jest w przyszłych pracach badawczych umożliwianie powtórnego podejścia do eksperymentu. Po drugie, w celu ograniczenia efektu habituacji konieczne może być zredukowanie poszczególnych list słów.

Do ograniczeń przeprowadzonego etapu badań zaliczyć można małą próbę badawczą oraz brak reprezentatywności pod względem równej liczby kobiet oraz mężczyzn czy też przewagę osób z wykształceniem wyższym.

Dyskusja ogólna

Na podstawie przeprowadzonych badań wnioskujemy, że przyszłe badania empiryczne z użyciem wykorzystanych przez nas list słów powinny uwzględniać jedynie połowę słów z poszczególnych kategorii. Powyższe zalecenie zostało sformułowane na podstawie dwóch argumentów.

Po pierwsze, do dalszych badań powinny zostać wykorzystane najlepsze słowa wchodzące w skład poszczególnych list wytypowane jako najbardziej podniecające. Nie rekomendujemy użycia do dalszych badań większej ilości słów niż połowa najlepszych słów z każdej listy, ponieważ poszczególne współczynniki zgodności sędziów kompetentnych nie pozwalają na czułe rozróżnianie między poszczególnymi słowami. Po redukcji poszczególnych list słów o połowę listy uzyskały akceptację sędziów kompetentnych.

Po drugie, na podstawie badań empirycznych przeprowadzonych na populacji nieklinicznej można założyć, że istotną rolę związaną z efektem eksperymentalnym odgrywa efekt habituacji. Wydaje się, że zaprezentowanie zbyt dużej liczby bodźców sprawia, że uczestnik badania uodparnia się na słowa prezentowane mu w czasie wykonywania zadania.

Zaprezentowane listy słów oraz rekomendacje badawcze obecnie nie dają podstawy do wykorzystania polskiej adaptacji zadania Stroopa w warunkach emocjonalnych do prowadzenia badań diagnostycznych. Brak możliwości wykorzystania zaprezentowanych list w diagnostyce seksuologicznej w obecnej chwili wynika z dwóch przesłanek. Po pierwsze, konieczne jest zweryfikowanie w późniejszych badaniach, czy stworzone listy słów różnicują osoby, które wykorzystwały seksualnie dzieci, oraz osoby z preferencjami pedofilnymi od osób z populacji nieklinicznej. Po drugie, niezbędne jest opracowanie wskaźników psychometrycznych oraz stworzenie odpowiednich danych normalizacyjnych dla poszczególnych grup sprawców przestępstw seksualnych.

Załącznik 1

Lista słów „neutralnych” oraz „kolorowych”

Załącznik nr 1

Lista słów „neutralnych”				Lista słów „kolorowych”			
Lp.	Słowo	Lp.	Słowo	Lp.	Słowo	Lp.	Słowo
1	tablica	16	pudełko	1	fioletowy	16	złoty
2	samochód	17	atlas	2	szary	17	srebrny
3	motocykl	18	sznurek	3	blady	18	jadeitowy
4	zdjęcia	19	pralka	4	różowy	19	topazowy
5	mikrofalówka	20	drukarka	5	pomarańczowy	20	szmaragdowy
6	toster	21	skaner	6	szkarłatny	21	fioletowy
7	czajnik	22	album	7	kasztanowy	22	szary
8	okulary	23	domofon	8	karmazynowy	23	blady
9	segregator	24	odkurzacz	9	biały	24	różowy
10	komputer	25	kabel	10	czarny	25	pomarańczowy
11	dysk			11	cyjan		
12	szalik			12	brązowy		
13	kamyk			13	cynowy		
14	projektor			14	brąz		
15	pierścień			15	fioletowy		

Piśmiennictwo

1. Nęcka E, Orzechowski J, Szymura B. *Psychologia poznawcza*. Warszawa: Wydawnictwo Szkoły Wyższej Psychologii Społecznej; 2008.
2. Christiansen P, Bloor JF. *Individualised but not general alcohol Stroop predicts alcohol use*. Drug Alcohol Depend. 2014; 134: 410–413.
3. Hill AB, Knowles TH. *Depression and the ‘emotional’ Stroop effect*. Pers. Individ. Differ. 1991; 12(5): 481–485.
4. Ciardha CÓ, Gormley M. *Using a pictorial-modified stroop task to explore the sexual interests of sexual offenders against children*. Sex. Abuse 2012; 24(2): 175–197.
5. Price SA, Beech AR, Mitchell I, Humphreys GW. *Measuring deviant sexual interest in adolescents using the emotional Stroop task*. Sex. Abuse 2014; 26(5): 450–471.
6. Smith P, Waterman M. *Processing bias for sexual material: The emotional Stroop and sexual offenders*. Sex. Abuse 2004; 16(2): 163–171.

7. Price SA, Hanson KR. *A modified Stroop task with sexual offenders: Replication of a study*. J. Sex. Aggress. 2007; 13(3): 203–216.
8. Leeuwen van ML, Baaren van RB, Chakhssi F, Loonen MG, Lippman M, Dijksterhuis A. *Assessment of implicit sexual associations in non-incarcerated pedophiles*. Arch. Sex. Behav. 2013; 42(8): 1501–1507.
9. Ayre C, Scally AJ. *Critical values for Lawshe's content validity ratio: revisiting the original methods of calculation*. Meas. Eval. Couns. Dev. 2014; 47(1): 79–86.

Adres: Wojciech Oronowicz-Jaškowiak
Instytut Psychiatrii i Neurologii
III Klinika Psychiatryczna
02-957 Warszawa, ul. Sobieskiego 9
e-mail: oronowicz@icloud.com

Otrzymano: 12.10.2018

Zrecenzowano: 24.06.2019

Otrzymano po poprawie: 19.07.2019

Przyjęto do druku: 14.08.2019