

Maciej Janowski

Uniwersytet Śląski

KONSTRUKCJA I WŁAŚCIWOŚCI PSYCHOMETRYCZNE TESTU WNIOSKOWANIA O EMOCJACH W SYTUACJI (TWES)

Umiejętność wnioskowania o emocjach wzbudzanych w określonych sytuacjach jest przedmiotem analiz w ramach badań nad inteligencją emocjonalną, empatią i teorią umysłu. W artykule zaprezentowano serię badań, prowadzących do powstania Testu Wnioskowania o Emocjach w Sytuacji i sprawdzających jego właściwości psychometryczne oraz trafność. Test złożony jest z 14 opisów sytuacji, a dla każdej z nich badany ma ocenić prawdopodobieństwo wystąpienia czterech wskazanych emocji. Do przeliczania punktacji wykorzystać można klucz konsensualny i targetowy – w obu przypadkach test charakteryzuje się dużą spójnością i rzetelnością (RMSEA poniżej 0,05; α -Cronbacha powyżej 0,7). Analizy korelacji z testami umiejętności emocjonalnych, umiejętności poznawczych, samooceną inteligencji emocjonalnej oraz kwestionariuszami osobowości potwierdzają trafność teoretyczną narzędzia przy zastosowaniu klucza targetowego. TWES może być stosowany w badaniach naukowych, obejmujących osoby w wieku od lat jedenastu.

Słowa kluczowe: umiejętności emocjonalne, rozumienie emocji, inteligencja emocjonalna

WPROWADZENIE

W ostatnich latach dużo uwagi w psychologii poświęcono umiejętnościom łączącym elementy funkcjonowania emocjonalnego i poznawczego. Jedną z nich jest rozumienie procesów emocjonalnych, opisywane w ramach inteligencji emocjonalnej (Mayer i Salovey, 1997; Mayer, 2000; zob. też pojęcie kompetencji emocjonalnej wg Saarni, 1999), empatii poznawczej (por.: Davis, 1980; Smith, 2006), teorii umysłu (Baron-Cohen, 2001) oraz wiedzy emocjonalnej (Izard, 2001).

W modelu Mayera i Saloveya (Mayer, Salovey, 1997; Mayer, 2000) rozumienie emocji, będące zastosowaniem wiedzy emocjonalnej (Mayer and Salovey, 1997), stanowi jeden z czterech głównych elementów inteligencji emocjonalnej, obok zdolności do postrzegania i wyrażania emocji, umiejętności uzyskiwania dostępu do własnych

przeżyć emocjonalnych lub ich generowania w sytuacjach gdy ułatwia to myślenie oraz zdolności do regulacji stanów emocjonalnych w sposób sprzyjający emocjonalnemu i intelektualnemu rozwojowi (Mayer i Salovey, 1997; Mayer, Caruso i Salovey, 1999). Rozumienie emocji bazuje na ich rozpoznawaniu i obejmuje w najprostszej formie zdolność do nazywania, etykietowania odczuć. Na wyższym poziomie zaawansowania oznacza zdolność do wyjaśnienia przyczyn emocji, ich znaczenia relacyjnego. Jeszcze bardziej złożona jest umiejętność dostrzegania możliwości jednoczesnego przeżywania różnych stanów emocjonalnych, a najwyższy poziom oznacza zdolność do rozpoznawania prawdopodobnych przemian, prowadzących od jednej emocji do drugiej, np. od złości do wstydu (Mayer, Salovey, 1997).

Rozumienie emocji obejmuje zatem między innymi zdolność do wnioskowania o cudzych

emocjach na podstawie analizy kontekstu sytuacyjnego, określaną czasami jako *emotion situation knowledge* (ESK; Fine, Izard i Trentacosta, 2006). Zdolność ta odgrywa ważną rolę w funkcjonowaniu interpersonalnym, ponieważ umożliwia rozpoznawanie cudzych emocji pomimo braku lub fałszowania ekspresji emocjonalnej (Blickle i in., 2009; Fine i in., 2006; Mayer i Salovey, 1997).

Dotychczasowe badania wskazują, że istnieją związki rozumienia emocji z akceptacją ze strony rówieśników, natężeniem przejawów niedostosowania społecznego (zob. przegląd w: Trentacosta, Izard, Mostow i Fine, 2006), wynikami w nauce (Hoerger, Chapman, Epstein i Duberstein, 2012; Rode i in., 2008), funkcją pełnioną w pracy i osiąganymi wynikami, szczerością, bystrością społeczną i wpływem interpersonalnym (Blickle i in., 2009), miarami afektu (Rode i in., 2008), ciepłem rodzicielskim i zadowoleniem z życia (Mayer, Caruso i Salovey, 1999) oraz kondycją psychiczną (zob. Berrocal i Extremera, 2008; Hansen, Lloyd i Stough, 2009; Lizeretti i Extremera, 2011). Trudno określić relacje z wymiarami osobowości, ze względu na brak powtarzalności uzyskiwanych korelacji z ekstrawertycznością, ugodowością i otwartością (zob.: Austin, Farrelly, Black i Moore, 2007; Blickle i in., 2009; Day i Carroll, 2004; Libbrecht i Lievens, 2012; Rode i in., 2008).

POMIAR ROZUMIENIA EMOCJI

Rozumienie emocji badane bywa za pomocą kwestionariuszy samoopisowych, analogicznych do kwestionariuszy osobowości albo poprzez zadania testowe, podobnie jak inteligencja.

Kwestionariusze i testy

W badaniach anglojęzycznych wykorzystuje się kilka kwestionariuszy, które przynajmniej częściowo odnoszą się do rozumienia emocji (Bar-On, 2006; Gignac, 2010; Salovey, Mayer, Goldman, Turvey i Palfai, 1995; Sharma, Gan-

gopadhyay, Austin i Mandal, 2013). W Polsce stosować można Popularny Kwestionariusz Inteligencji Emocjonalnej (Jaworowska i Matczak, 2005), zawierający podskale rozumienia emocji własnych i cudzych. Weryfikowanie umiejętności emocjonalnych na drodze samooceny budzi jednak duże zastrzeżenia, między innymi ze względu na nietrafność obrazu swoich zdolności (Brackett, Rivers, Shiffman, Lerner i Salovey, 2006; Zeidner, Shanizinovich, Matthews i Roberts, 2005) oraz możliwość intencjonalnego fałszowania odpowiedzi (zob.: Grubb III i McDaniel, 2007; Whitman, Van Rooy, Viswesvaran i Alonso, 2008). Mayer i współpracownicy (np. Mayer, Salovey i Caruso, 2000; Mayer, Salovey, Caruso i Sitarenios, 2003) postulują więc wykorzystywanie zadań testowych, w których wynik zależny jest od poziomu wykonania, a nie samoopisu.

Najczęściej wykorzystywanym testem umiejętności jest Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT; Brackett i Salovey, 2006; Mayer i in., 2003). Niestety jego podskale diagnozujące rozumienie emocji mają niską rzetelność (np. Austin, 2010; MacCann, Matthews, Zeidner i Roberts, 2003; Maul, 2011; O'Connor i Little, 2003; Roberts i in., 2006). Strukturę podobną do MSCEIT ma opracowany przez Śmieję i Orzechowskiego Test Inteligencji Emocjonalnej (Śmieja, Orzechowski i Stolarski, 2014; Śmieja, Orzechowski i Beauvale, 2007), zawierający podskale rozumienia emocji, w której osoba badana ma określić prawdopodobieństwo odczuć osób w opisywanych sytuacjach. Relacjonowana rzetelność tej części testu wynosi 0,68 (Śmieja i in., 2007) i 0,69 (Śmieja i in., 2014).

W literaturze opisanych zostało także kilka testów, badających wyłącznie rozumienie emocji. Situational Test of Emotional Understanding (STEU; MacCann, 2006; MacCann i Roberts, 2008) i Test of Emotional Intelligence (TEMINT; Blickle i in., 2009; Blickle, Momm, Liu, Witzki i Steinmayr, 2011) zawierają opisy sytuacji, a osoba badana ma oszacować prawdopodobieństwo lub natężenie wskazanych reakcji emocjonalnych. Nieco inny charakter ma Test

Rozumienia Emocji (TRE; Matczak i Piekarska, 2011) ponieważ składa się z pięciu podtestów, badających różne aspekty tej umiejętności. W kilku badaniach uzyskano zadowalającą wartość współczynnika rzetelności STEU (Libbrecht i Lievens, 2012; MacCann, Pearce i Roberts, 2011; MacCann i Roberts, 2008) przy najwyższej wartości $\alpha=0,75$ (MacCann, 2006), ale w części badań współczynnik Cronbacha był niski i nie osiągnął wartości 0,7 (Austin, 2009, 2010; MacCann, 2010; MacCann, 2006; MacCann i Roberts, 2008). Wyższą rzetelnością charakteryzuje się TEMINT, dla którego α -Cronbacha wynosi od 0,71 do 0,82 (Blickle i in., 2009, 2011) oraz TRE, dla którego autorki podają wartość od 0,78 do 0,84 w zależności od grupy (Matczak i Piekarska, 2011). Powyższe dane wskazują, że bardziej rzetelny pomiar ma miejsce przy zastosowaniu testów jednowymiarowych, przygotowanych z myślą o diagnozie konkretnej części rozumienia emocji.

Określanie poprawności odpowiedzi w testach

W budowaniu testu umiejętności emocjonalnych jednym z głównych problemów jest zawsze określenie właściwego klucza odpowiedzi (Petrides i Furnham, 2001). W przypadku inteligencji poznawczej zazwyczaj zadania budowane są tak, aby właściwa była jedna odpowiedź, którą można wyłonić dzięki logicznemu rozumowaniu, informacjom słownikowym itp. W przypadku reakcji emocjonalnych jest to raczej nieosiągalne, ze względu na różnice indywidualne i dużą rolę kontekstu w budowaniu reakcji afektywnej. Wnioskowanie o procesach emocjonalnych ma raczej charakter probabilistyczny, a nie dedukcyjny. Z trudnością tą twórcy testów najczęściej próbują poradzić sobie przez zastosowanie punktacji opartej o klucz konsensualny (MSCEIT) lub klucz ekspercki (MSCEIT, TRE, TIE).

W kluczu eksperckim sędziowie kompetentni, najczęściej psychologowie, wskazują odpowiedzi najlepsze. W kluczu konsensualnym punktacja uzależniona jest od tego, jak test wypełniała

wcześniej duża grupa osób badanych – im większa zbieżność z odpowiedziami większości, tym więcej punktów zyskuje badany. Podejście takie bywa krytykowane, ponieważ uniemożliwia wyodrębnienie zadań najtrudniejszych, rozwiązywanych tylko przez osoby najbardziej inteligentne (Matthews, Zeinder i Roberts, 2004).

Odmienne metody ustalania klucza wykorzystano przy tworzeniu TEMINT i STEU. W TEMINT punktacja zbudowana jest na danych osób (*target persons*), które miały scharakteryzować własne reakcje emocjonalne w opisywanych sytuacjach (Blickle i in., 2011). Im większa zgodność odpowiedzi osoby badanej z odpowiedziami *target persons*, tym lepsze jej rozumienie emocji. STEU skonstruowany został w taki sposób, aby prezentowane w nim sytuacje były zgodne z charakterystyką oceny poznawczej, prowadzącej do określonych emocji w koncepcji Roseman (Roseman, Spindel i Jose, 1990), a poprawne odpowiedzi wyznaczone zostały na podstawie tej samej teorii (MacCann i Roberts, 2008).

ZAŁOŻENIA PRZYJĘTE PRZY TWORZENIU TWES

Podstawą pracy nad testem było przyjęcie założenia, że rozumienie emocji jest związane z wiedzą o emocjach i jej wykorzystaniem (Blickle i in., 2009; Matczak i Piekarska, 2011; Mayer i Salovey, 1997), dlatego do jego badania lepiej stosować zestaw zadań, niż kwestionariusz samooceny (por.: Mayer i in., 1999, 2000, 2003). Jak wskazują dane zaprezentowane powyżej, w testach wielowymiarowych trudno osiągnąć zadowalającą rzetelność poszczególnych podskal, zatem przedmiotem analiz będzie tylko jedna ze zdolności, składających się na rozumienie procesów emocjonalnych, mianowicie umiejętność wnioskowania o tym, co odczuwa druga osoba, znajdująca się określonej sytuacji. Innymi słowy test ma badać zdolność do wskazania na podstawie kontekstu sytuacyjnego tych emocji, które są prawdopodobną reakcją opisywanej osoby.

Praca nad nowym testem została rozpoczęta z myślą o stworzeniu narzędzia, które umożliwiłoby badanie zmian rozwojowych w obrębie rozumienia emocji od dzieciństwa do dorosłości. Dostępny w języku polskim Test Rozumienia Emocji przeznaczony jest do badania osób w wieku od 15 lat (Matczak i Piekarska, 2011), a w literaturze przedmiotu brak systematycznych badań nad rozumieniem procesów emocjonalnych dzieci i młodzieży. Ze względu na mniejszą dojrzałość poznawczą dzieci zarówno formę zadań, jak i sposób udzielania odpowiedzi należało jak najbardziej uprościć. Jako minimalny wiek stosowania nowego testu przyjęto 11 lat. Jak wskazują badania, dzieci w tym wieku są w stanie zrozumieć, że jednocześnie można przeżywać sprzeczne emocje wobec tego samego obiektu (Burkitt i Watling, 2015; Harter i Buddin, 1987; Reissland, 1985), także takie, które wymagają reprezentacji stanów umysłowych innych osób, np. poczucie winy lub zażenowanie (Burnett, Thompson, Bird i Blakemore, 2011) oraz uwzględniać specyficzne informacje o osobie przeżywającej emocje (Saarni, 1999).

Reakcje afektywne wykazują duże zróżnicowanie indywidualne i są zależne nie tylko od kontekstu sytuacyjnego, ale także od oceny poznawczej, osadzonej w celach, wartościach, sieci skojarzeń podmiotu. Klucz punktowania odpowiedzi postanowiono zatem zbudować na podstawie danych z samooceny reakcji emocjonalnych, uzyskanych w badaniu wstępnym, podobnie jak uczynili to autorzy testu TEMINT. Ze względu na wagę społecznych skryptów emocjonalnych, uwzględniony zostanie dodatkowo klucz konsensualny, stworzony na podstawie odpowiedzi dużej grupy osób badanych w teście docelowym.

ETAPY TWORZENIA TESTU

Dobór pozycji testu

W pierwszym etapie pracy na testem, na podstawie materiałów z opracowania MacCann (2006) oraz własnych pomysłów, przygotowano

kwestionariusz zawierający 42 opisy sytuacji, oraz listę 11 emocji. Arkusz przekazano 508 osobom w wieku od 10 do 82 lat, z prośbą o wskazanie emocji, które prawdopodobnie przeżywałyby w prezentowanych wydarzeniach. Na podstawie uzyskanych wyników wyodrębniono 17 historyjek o najbardziej zróżnicowanych reakcjach oraz przyporządkowano każdej z nich 4 emocje. W ten sposób powstał kwestionariusz preTWES, który przekazano kolejnej grupie 476 osób w wieku od 18 do 74 lat ($M_{\text{wiek}}=28,5$; $SD_{\text{wiek}}=12$), w tym 300 kobiet i 176 mężczyzn. Badani proszeni byli o zaznaczenie, z możliwością dokonania wyborów wielokrotnych, którymi emocjami zareagowałoby, gdyby przydarzyło im się coś podobnego. W odniesieniu do kilku opisów zastosowana została nieco zmodyfikowana konstrukcja: najpierw pytano o odczuwanie jednej, wskazanej, najmocniejszej emocji, a następnie proszono o zaznaczenie pozostałych czterech, jak powyżej.

Przeformułowanie pozycji testowych i przygotowanie klucza odpowiedzi

Analiza wyników preTWES doprowadziła do odrzucenia trzech pozycji ze względu na małą wariancję odpowiedzi, a jednocześnie ujawniła różnice płciowe w reakcjach na cztery sytuacje bodźcowe. Opisy wydarzeń z preTWES zostały przeformułowane tak, aby dotyczyły osoby trzeciej, przy czym trzy pozycje, w których odnotowano różnice płciowe, zostały zdublowane – raz bohaterem była kobieta, za drugim razem mężczyzna. Po dołączeniu do nich zestawu emocji z preTWES powstała wstępna wersja Testu Wnioskowania o Emocjach w Sytuacji (TWES18), zawierająca 18 opisów sytuacji. Osoba badana miała ocenić, czy każda z czterech emocji jest prawdopodobną reakcją na opisane wydarzenie. Dodatkowo, na początku arkusza, wprowadzone zostały dwie pozycje buforowe. Po przeprowadzeniu wstępnego badania na grupie 100 osób odrzucono dwie pozycje o niskiej rzetelności, a spośród itemów zdublowanych wybrano te, które dawały wyniki bardziej spójne z pozostałymi.

W efekcie powstał arkusz TWES, zawierający dwanaście opisów sytuacji testowych i dwa buforowe. Pytania buforowe nie są punktowane i zawierają listę emocji mało prawdopodobnych w opisanych sytuacjach. Ich wprowadzenie wynika z obserwacji, że wiele osób ma tendencję do zbyt częstego wykorzystywania odpowiedzi *TAK / RACZEJ TAK*. Pozycje buforowe mają zatem na celu uzmysłowienie badanemu, że zaznaczanie tej odpowiedzi w wielu przypadkach nie jest konieczne. Przy każdym opisie umieszczono cztery emocje, a osoba badana miała oszacować w skali trzypunktowej (*TAK/RACZEJ TAK – MOŻE – NIE/RACZEJ NIE*), czy każda ze wskazanych czterech emocji może pojawić się u bohatera opisywanego wydarzenia. Przykładową pozycję z TWES prezentuje rycina 1. W części badań zastosowano zmodyfikowaną wersję testu, w której przy każdej sytuacji należało dodatkowo wskazać emocję najmocniejszą.

Wyniki badania arkuszem preTWES posłużyły także do konstrukcji klucza testu TWES, określanego jako targetowy (TWES_t). Początkowo przyjęto, że jeśli nie więcej niż 1/3 badanych w preTWES wskazała daną emocję jako prawdopodobną, to w TWES_t przyznawane są 3 punkty za odpowiedź NIE, jeżeli więcej niż 2/3, to na 3 punkty zaliczana jest odpowiedź TAK, a w pozostałych przypadkach 3 punkty uzyskuje się za odpowiedź MOŻE. Za odpowiedzi nieoptymalne można uzyskać 0, 1 lub 2 punkty, w zależności od odsetka osób wskazujących daną emocję w preTWES. Kryterium poprawności odpowiedzi w TWES_t jest zatem związane z proporcją osób, które w badaniu

wstępnym określiły daną emocję jako prawdopodobną własną reakcją w opisanej sytuacji. Wysoką punktację uzyskuje się za wskazanie dużego prawdopodobieństwa tych emocji, które w preteście zaznaczyło dużo osób, albo za określenie jako mało prawdopodobnych tych emocji, które w preteście były wskazywane bardzo rzadko. Przegląd wyników pokazał jednak, że odpowiedzi z wysoko punktowaną odpowiedzią NIE jest aż 25 i łącznie można za nie uzyskać około 70% punktów, w związku z tym zmniejszono maksymalną punktację za tę odpowiedź do 2. Wprowadzony został też dodatkowy wskaźnik TWES_{tP}, w którym wynik jest pomniejszony o iloczyn sumy punktów za odpowiedzi NIE w 25 wspomnianych pozycjach i proporcję odpowiedzi NIE przy pięciu emocjach, w których należało zaznaczyć odpowiedź TAK. Wskaźnik ten obliczany jest według wzoru:

$$TWES_{tP} = TWES_t - SPN \times \frac{LON}{5}$$

gdzie

SPN oznacza sumę punktów za odpowiedzi NIE w 25 pozycjach, w których za NIE otrzymuje się 2 punkty;

LON oznacza liczbę odpowiedzi NIE przy pięciu emocjach, które co najmniej 2/3 badanych preTWES wskazało jako prawdopodobne reakcje na opisywane sytuacje.

Drugi klucz, określaný jako konsensualny (TWES_k), powstał na podstawie odpowiedzi osób wypełniających wcześniej TWES. Spośród 2879 badanych, za pomocą doboru losowego warstwowego (ze względu na wiek i płeć)

		TAK/ RACZEJ TAK	MOŻE	RACZEJ NIE/ NIE
Pewna osoba jedzie autobusem. Jakiś mężczyzna co chwilę kaszle nad nią, nie odwracając się w inną stronę, ani nie zakrywając ust. Co czuje ta osoba?	nadzieja			
	strach			
	złość			
	przykrość			

Ryc 1. Przykładowe zadanie z TWES

dobrano grupę referencyjną, składającą się z 526 osób w wieku od 20 do 79 lat ($M_{\text{wiek}}=42,3$; $SD_{\text{wiek}}=14,5$) w równych proporcjach płciowych. Szerszą charakterystykę wieku w tej grupie zawiera tabela 1. W TWESk liczba punktów uzależniona jest od tego, ile osób w grupie referencyjnej zaznaczyło daną odpowiedź (por. Mayer i in., 2003), w związku z tym maksymalna punktacja jest różna dla poszczególnych odpowiedzi i waha się w granicach od 4 do 9.

Klucz konsensualny odwołuje się do wyobrażeń większości osób na temat reakcji afektywnych innych ludzi. Wysoka punktacja świadczy o tym, że wizja funkcjonowania emocjonalnego, prezentowana przez osobę badaną, jest zgodna z obrazem posiadanym przez większość ludzi w naszym kręgu kulturowym, ze swego rodzaju standardem kulturowym lub „psychologią potoczną” w zakresie funkcjonowania emocjonalnego człowieka. Mówiąc inaczej, badani uzyskujący wysoką punktację w tym kluczu przewidują reakcje emocjonalne tak, jak większość ludzi. Można ją interpretować jako element dostosowania społecznego lub przejaw stereotypowego myślenia o funkcjonowaniu człowieka. Niskie wyniki oznaczają niestandardowe spostrzeżenie reakcji emocjonalnych, wynikające bądź z braku rozeznania reakcji emocjonalnych, bądź z bardzo wysokiego poziomu rozumienia związków między sytuacją a odczuciem. Druga możliwość wynika z założenia, że analizowana umiejętność ma rozkład normalny, a osoby najlepiej wnioskujące o emocjach stanowią stosunkowo niewielką grupę, odpowiadającą inaczej niż większość, dlatego nie uzyskują wysokiej punktacji.

Wysoka punktacja w kluczu targetowym jest związana z umiejętnością wskazania emocji, które u większości osób rozwiną się w określonym kontekście sytuacyjnym, ale także ze zdolnością do określenia, które emocje raczej nie pojawią się w tej sytuacji. Niski wynik może być wskaźnikiem braku rozeznania w reakcjach emocjonalnych, a jedną z jej form może być skłonność do przeceniania, przeszacowania prawdopodobieństwa różnych odczuć. W konsekwencji przyjąć można, że badani o najwyższych wynikach TWESk przewidują w różnych sytuacjach o wiele szerszą gamę emocji, niż osoby, które otrzymują dużo punktów w TWESk.

SPRAWDZENIE WŁAŚCIWOŚCI PSYCHOMETRYCZNYCH TWES

Badanie pierwsze miało na celu określenie struktury testu i jego zgodności wewnętrznej. Z założenia TWES miał mieć jednolitą strukturę, zatem analizy czynnikowe były prowadzone przede wszystkim w kierunku jej potwierdzenia. Sprawdzono także związki z wiekiem i płcią.

Choć badań nad rozwojem rozumienia emocji jest niewiele (np.: Davis i Humphrey, 2014; Fine i in., 2006), przyjąć można, że umiejętność ta rozwija się w okresie dzieciństwa i dorastania, podobnie jak inne umiejętności poznawcze. Badania Tonks i współpracowników (Tonks, Williams, Frampton, Yates i Slater, 2007) wskazują jednak, że zależność między umiejętnościami emocjonalnymi a wiekiem może nie mieć charakteru prostoliniowego.

Tabela 1. Charakterystyka wieku osób, których wyniki zostały wykorzystane do stworzenia klucza konsensualnego

Wiek	N	M	SD	Skośność	Kurtoza
20–29	116	23,53	2,63	0,32	-0,84
30–39	116	33,53	2,98	0,43	-1,09
40–49	116	45,15	2,64	-0,34	-0,84
50–59	116	53,75	2,81	0,22	-1,11
60+	62	67,45	5,40	0,35	-0,86

W wielu badaniach kobiety uzyskują wyższe wyniki w testach umiejętności emocjonalnych niż mężczyźni (np.: Matczak i Piekarska, 2011; Salguero, Extremera i Fernández-Berrocal, 2012; Siegling, Saklofske, Vesely i Nordstokke, 2012; Williams, Daley, Burnside i Hammond-Rowley, 2009), zatem podobnej prawidłowości można spodziewać się w odniesieniu do TWES.

Wykorzystane narzędzie i procedura badawcza

Badanie przeprowadzono za pomocą TWES. Do badania dzieci i młodzieży arkusz został nieco zmodyfikowany: w instrukcji wykorzystano bezpośredni zwrot do respondenta, dwie sytuacje, które w wersji dla dorosłych dotyczyły pracy, zostały osadzone w kontekście szkolnym, a sytuacja wygranej w totolotku została zamieniona na sytuację wygranej w loterii szkolnej. Badanie miało charakter indywidualny lub grupowy, a uczestników pozyskiwano metodą kuli śnieżnej. Czas wypełniania testu wahał się w granicach 5–20 minut.

Opis grupy badawczej

Do analiz wykorzystano wyniki kilku grup, zróżnicowanych pod względem wykształcenia i zawodu. Łącznie uwzględniono dane od 2657 osób w wieku od 11 do 65 lat. Liczebność poszczególnych grup wiekowych z uwzględnieniem płci zawiera tabela 2.

Tabela 2. Liczba osób biorących udział w badaniu 1 z podziałem według płci i grupy wiekowej

Grupa wiekowa	Łącznie ^a	Kobiety	Mężczyźni
8–12	99	44	54
13–15	115	61	54
16–18	606	263	214
19–25	485	343	136
26–40	386	270	107
41–65	415	266	144
Razem ^b	2154	1273	726

^a Uwzględniono wyniki osób, które nie wskazały swojej płci.

^b Uwzględniono wyniki osób w wieku powyżej 65 lat.

Prezentacja wyników

Analizy właściwości pozycji testowych przeprowadzono każdorazowo z wykorzystaniem kluczy TWES_t i TWES_k. Obliczenia we wszystkich badaniach przeprowadzone zostały za pomocą oprogramowania STATISTICA 10.

Struktura testu

Pierwszym krokiem było sprawdzenie struktury TWES poprzez eksploracyjną analizę czynnikową metodą składowych głównych z rotacją Varimax. Statystyki KMO (0,80 dla TWES_t; 0,81 dla TWES_k) oraz test sferyczności Bartletta (dla TWES_t: $\chi^2=11982$, $df=803$, $p<0,001$; dla TWES_k: $\chi^2=13768$, $df=903$; $p<0,001$) wskazują na możliwość stosowania analizy czynnikowej. Na podstawie wykresu osypiska sprawdzano modele trzy- i pięcioczynnikowy dla TWES_t oraz czteroczynnikowy dla TWES_k. We wszystkich przypadkach wartość ładunków jest niska (maksymalnie 0,56; 0,69 dla TWES_t oraz 0,65 dla TWES_k), podobnie jak procent wyjaśnianej przez nie łącznie ogólnej wariancji (17,5%; 23,9% dla TWES_t i 23,6% dla TWES_k), zatem wyodrębnianie podskal TWES nie ma podstaw.

Spójność testu sprawdzono za pomocą współczynnika α -Cronbacha. Wstępne analizy doprowadziły do usunięcia z obliczeń pięciu pozycji dla obu kluczy. Uzyskane wyniki po ich usunięciu zawiera tabela 3.

Po wyeliminowaniu pozycji słabo korelujących z wynikiem całkowitym, sprawdzona została hipoteza o jednoczynnikowej strukturze TWES za pomocą confirmacyjnej analizy czynnikowej z zastosowaniem funkcji zbieżności GLS→ML. W analizie wykorzystano wyniki poszczególnych odpowiedzi (emocji). Test kurtozy Mardii wskazuje, że dane mogą być użyte do budowania modelu: $Kappa_{TWES_t}=0,034$; $Kappa_{TWES_k}=0,097$. Przy zastosowaniu obu kluczy wskaźniki dobroci dopasowania są zadowalające (por: Sagan, 2003): dla TWES_t – $\chi^2=4283$; $df=860$; $p<0,001$; $\chi^2/df=4,98$; RMSEA=0,049; AGFI=0,888, dla TWES_k – $\chi^2=4992$, $df=860$; $p<0,001$; $\chi^2/df=5,80$; RMSEA=0,052; AGFI=0,886.

Tabela 3. Wskaźniki α -Cronbacha dla testu TWES w poszczególnych grupach wiekowych (po usunięciu 5 pozycji dla TWES*t* i 5 dla TWES*k* obniżających rzetelność)

Grupa wiekowa	TWES <i>t</i>			TWES <i>k</i>		
	C	K	M	C	K	M
11–12	0,73	0,67	0,75	0,80	0,82	0,72
13–15	0,74	0,75	0,72	0,71	0,56	0,78
16–18	0,72	0,73	0,71	0,77	0,74	0,82
19–25	0,76	0,76	0,76	0,74	0,71	0,74
25–40	0,79	0,80	0,76	0,81	0,81	0,80
41–65	0,74	0,74	0,73	0,76	0,76	0,77
Całość	0,75	0,76	0,74	0,78	0,77	0,77

C – całość; K – kobiety; M – mężczyźni.

Wyniki obliczane za pomocą odrębnych kluczy korelują ze sobą ujemnie (dla TWES*t* i TWES*k* współczynnik korelacji Spearmana $r_s = -0,68$; dla TWES*tP* i TWES*k* $r_s = -0,58$; w obu przypadkach $p < 0,001$), ale związki między TWES*tP* i TWES*k* nie są prostoliniowe: w przedziale do 160 punktów w TWES*k* $r_s = 0,38$ ($p < 0,001$; $N = 98$), a powyżej 160 punktów $r_s = -0,59$ ($p < 0,001$; $N = 2519$).

Normalność rozkładu wyników i trudność zadań

Wyniki testów normalności rozkładu, wraz z podstawowymi statystykami opisowymi, zamieszczone zostały w tabeli 4.

Na podstawie testu Kołmogorowa-Smirnowa, wielkości skośności i kurtozy oraz ich błędów standardowych ($SES = 0,05$; $SEK = 0,1$) nie można

przyjąć założenia o rozkładzie normalnym wyników TWES. TWES*t* i TWES*tP* są nieco prawoskośne, a wyniki TWES*k* mają rozkład lewoskośny i leptokurtyczny. Z tego względu w obliczeniach stosowane będą statystyki nieparametryczne.

Brak normalności rozkładu związany jest poziomem trudności zadań. W TWES*k* poziom ten, obliczany jako iloraz średniej i maksymalnej punktacji dla danej pozycji, waha się między 0,72 a 0,93 przy średniej 0,83. Przy kluczu konsensualnym TWES jest zatem stosunkowo łatwy: 81% osób uzyskało wyniki w granicach 183–234 punkty, podczas gdy wynik najwyższy wyniósł 244 punkty. Dla TWES*t* poziom trudności mieści się w przedziale 0,12–0,63, przy średniej 0,35. Informacje na temat trudności wszystkich zadań zawierają materiały dodatkowe (informacja na końcu artykułu).

Tabela 4. Statystyki opisowe TWES oraz wyniki testu normalności rozkładu

Skala	Średnia	Mediana	Minimum ^a	Maksimum ^a	SD	Skośność	Kurtoza	d	Poziom istotności dla d
TWES <i>t</i>	35,88	35	4 / 0	80 / 115	12,54	0,31	-0,15	0,05	<0,01
TWES <i>tP</i>	34,40	33	4 / 0	80 / 115	12,21	0,42	0,04	0,06	<0,01
TWES <i>k</i>	204,9	209	75 / 70	244 / 247	22,7	-1,18	2,14	0,08	<0,01

d – wartość statystyki Kołmogorowa-Smirnowa z poprawką Lillieforsa.

^a Wartość obserwowana / wartość teoretyczna.

Wyniki testu a wiek i płeć badanych

Wskaźniki korelacji między wiekiem a wynikami TWES zawiera tabela 5.

U dzieci i młodzieży wyniki TWES korelują z wiekiem dodatnio przy zastosowaniu klucza konsensualnego, ale przy uwzględnieniu płci efekt ten ujawnia się tylko w grupie dziewcząt. U chłopców korelacje istotne odnotowano w przypadku klucza targetowego. W dorosłości wyniki obliczane kluczem targetowym są pozytywnie związane z wiekiem u kobiet, a TWESk w tej grupie koreluje z wiekiem ujemnie. Wszystkie relacjonowane korelacje mają jednak niskie wartości.

Porównania grup płciowych wskazują, że kobiety uzyskały istotnie wyższe wyniki w TWESk ($Z_{TWESk}=5,94$; $p<0,001$; Z odpowiada wartości U Manna-Whitneya) natomiast w kluczu targetowym nieco wyższą punktację uzyskali mężczyźni ($Z_{TWESt}=-2,71$; $p<0,05$; $Z_{TWEStP}=-1,81$; $p<0,05$). Przy analizie z uwzględnieniem wieku różnice istotne pojawiły się w grupach wiekowych 16–18 lat i 19–24 lata – kobiety uzyskały więcej punktów w TWESk, mężczyźni w kluczu targetowym. Pełne zestawienie wyników można znaleźć w materiałach dodatkowych.

Dyskusja wyników

Wyniki TWES nie mają rozkładu normalnego bez względu na zastosowany klucz, przy czym odchylenie od normalności jest wyraźne dla wyników TWESk – duża liczba badanych uzyskała wysoką punktację.

Wyniki analizy czynnikowej nie wskazują na wewnętrzne zróżnicowanie testu, wskaźniki rzetelności α -Cronbacha są zadowalające do badań naukowych we wszystkich przedziałach wiekowych przy obu kluczach. Wyjątek stanowią rezultaty dziewcząt w wieku gimnazjalnym, dla których rzetelność jest niska. Test można stosować w badaniu uczniów ze starszych klas szkoły podstawowej.

Łączna analiza wyników korelacji TWES z wiekiem oraz porównań między grupami płciowymi prowadzi do wniosku, że w okresie dorastania u dziewcząt stopniowo rozwija się skłonność do wnioskowania o szerokiej gamie odczuć, przeżywanych w różnych sytuacjach, zgodnie z funkcjonującymi w społeczeństwie przekonania na temat emocji. Wnioskowanie takie może być częściowo nietrafne i w okresie dorosłości następuje jego korygowanie, dzięki czemu kobiety osiągają poziom zbliżony do prezentowanego przez mężczyzn mniej więcej w czwartej dekadzie życia. Chłopcy w mniejszym stopniu przejmują kulturowe standardy funkcjonowania emocjonalnego i wraz z rozwojem coraz trafniej potrafią nie tylko rozpoznać te emocje, które prawdopodobnie pojawią się w danej sytuacji, ale także wskazać, że część odczuć jest mało prawdopodobną reakcją.

Zaskakujące mogą się wydawać korelacje między wynikami obliczonymi za pomocą odrębnych kluczy. Wysoka punktacja w kluczu targetowym związana jest w dużym stopniu z umiejętnością wskazania tych emocji, które nie powstają

Tabela 5. Korelacje TWES z wiekiem (rs Spearmana)

Skala	Wiek 11–18 lat			Wiek 19–65 lat			Wiek 11–65 lat		
	K	M	C	K	M	C	K	M	C
TWES _t	0,00	0,14*	0,04	0,15***	0,07	0,13***	0,12***	0,10**	0,10***
TWES _{tP}	0,02	0,15**	0,06	0,10**	0,07	0,10***	0,10***	0,11**	0,10***
TWES _k	0,13*	0,03	0,11***	-0,24***	-0,10*	-0,21***	-0,04	0,04	0,00

K – kobiety; M – mężczyźni; C – cała grupa, bez różnicowania płci.

*** $p<0,001$; ** $p<0,01$; * $p<0,05$.

w danej sytuacji, czyli z zaznaczeniem odpowiedzi NIE. Osoby badane mają natomiast skłonność do częstego zaznaczania w TWES opcji TAK i MOŻE, dlatego klucz konsensualny, budowany na ich wynikach, „premiuje” takie odpowiedzi. W efekcie TWES*t* i TWES*k* korelują ujemnie. Fakt ten można też wiązać ze sposobem punktowania odpowiedzi w TWES*k*, potencjalnie niekorzystnym dla osób o najwyższym poziomie umiejętności. Jeśli przyjąć, że zdolność do wnioskowania o emocjach na podstawie wskaźników sytuacyjnych ma rozkład normalny w populacji, to większość badanych prezentuje umiarkowany jej poziom, jednocześnie uzyskując wysokie wyniki w kluczu konsensualnym. Wyższy poziom umiejętności prezentowany jest przez mniejszą liczbę osób, dlatego uzyskują one wyniki niższe w TWES*k* (nt. problematyczności punktacji konsensualnej w ocenie małych podgrup patrz: Barchard i Russell, 2006).

Być może rozbieżność między wynikami w kluczu targetowym i konsensualnym związane są także z różnicami w wykorzystywanych skryptach emocji, czyli jednostkach poznawczych, opisujących w sposób uporządkowany czynniki wzbudzające, zmiany w organizmie, odczucia, formy ekspresji, zachowania i etykiety słowne właściwe dla danej emocji (zob. Fehr i Russell, 1984; Widen i Russell, 2011). TWES*t* związany jest z wykorzystaniem skryptów indywidualnych, które mogą różnić się od stanowiących podstawę dla TWES*k* skryptów dotyczących anonimowych członków grupy (zob. Janowski, 2015), co skutkuje odmienną punktacją w obu kluczach. Teza ta wymaga jednak potwierdzenia empirycznego.

STABILNOŚĆ BEZWZGLĘDNA I WIARYGODNOŚĆ TESTU

Do oszacowania rzetelności testu, prócz wspomnianego powyżej wskaźnika α -Cronbacha, wykorzystano także estymację stabilności bezwzględnej i wiarygodności (zob.: Brzeziński, 1999).

W badaniu określającym wiarygodność udział wzięło 95 studentów pierwszego roku pedagogiki (średnia wieku 19,4; SD=0,74). Po wypełnieniu TWES uczestnicy byli informowani, że pomiar musi być powtórzony i ponownie wypełniali arkusz. W drugim pomiarze TWES*t* i TWES*tP* badani uzyskali istotnie wyższe wyniki, niż w pomiarze pierwszym ($p < 0,001$ w teście znaków). Dla TWES*k* wyniki są takie same.

Badania służące określeniu stabilności odbyły się w odstępie trzech tygodni. Udział w nim wzięli studenci pierwszego roku psychologii, teologii i uczniowie liceum ogólnokształcącego. Łącznie udało się zgromadzić wyniki 37 osób ($M_{\text{wiek}}=17,8$; $SD_{\text{wiek}}=2,11$). Przy zastosowaniu testu znaków nie odnotowano różnic między wynikami w obu pomiarach.

Wyniki analiz korelacyjnych zawiera tabela 6. Fakt, że punktacja drugiej tury w badaniu wiarygodności jest wyższa niż pierwszej, może wskazywać na podatność testu na korygowanie odpowiedzi wskutek uczenia się. Brak różnic w przypadku stabilności można interpretować jako przejaw zanikania efektu uczenia się, ale może też być artefaktem, wynikającym z mniejszej liczebności grupy badawczej.

Tabela 6. Wskaźniki korelacji Spearmana między dwoma pomiarami TWES

Wskaźnik rzetelności	TWES <i>t</i>	TWES <i>tP</i>	TWES <i>k</i>
Pomiar natychmiastowy (N=95)	0,80***	0,80***	0,72***
Wiarygodność estymowana	0,84	0,84	0,83
Pomiar w odstępie 3 tygodni (N=37)	0,59***	0,54***	0,47**
Rzetelność estymowana	0,78	0,73	0,70

Wiarygodność i rzetelność estymowana – współczynnik rzetelności obliczony z uwzględnieniem różnic między wielkością wariancji w dwóch grupach.

*** $p < 0,001$; ** $p < 0,01$.

Wiarygodność testu jest zadowalająca, natomiast stabilność czasowa w odstępie trzytygodniowym jest umiarkowana. Należy jednak

zwrócić uwagę, że grupa, która brała udział w weryfikacji stabilności jest bardzo homogeniczna ($SD_{TWEST}=9,11$; $SD_{TWESTP}=9,32$; $SD_{TWESK}=17,22$). Z tego względu dodatkowo oszacowana została wartość współczynnika rzetelności z uwzględnieniem różnic między wielkością wariancji w dwóch grupach (wzór 5.27 w: Magnusson, 1991). W obliczeniach wykorzystano wartości odchyłeń standardowych z pierwszego pomiaru w badaniu wiarygodności oraz z tabeli 4. Tak oszacowane współczynniki mają zadowalające wartości.

TRAFNOŚĆ TESTU

Wyniki testów rozumienia zjawisk emocjonalnych powinny korelować z wynikami innych testów sprawdzającymi tę umiejętność, a w mniejszym stopniu ze wskaźnikami inteligencji poznawczej i rezultatami testów innych umiejętności emocjonalnych, na przykład rozpoznawania emocji. Należy też spodziewać się związków TWES z kwestionariuszami inteligencji emocjonalnej i empatii, przy braku istotnych korelacji z wymiarami osobowości (Petrides i Furnham, 2001).

Grupa badawcza i wykorzystane narzędzia

W celu sprawdzenia trafności TWES przeprowadzono serię badań z zaangażowaniem odrębnych grup. W określeniu korelacji z testami inteligencji poznawczej, testami umiejętności emocjonalnych oraz samooceną empatii wykorzystane zostały wyniki uczniów klas I–III szkół ponadgimnazjalnych oraz klasy III gimnazjum z terenu województwa śląskiego. W badaniach samooceny inteligencji emocjonalnej, empatii oraz osobowości udział wzięły osoby pozyskiwane metodą kuli śnieżnej. Informacje o wieku badanych zawierają tabele 7 i 8.

Do weryfikacji trafności TWES wykorzystano następujące narzędzia:

- Test Rozumienia Emocji (TRE; Matczak i Piekarska, 2011);

- test rozpoznawania emocji z głosu z zestawu Diagnostic Analysis of Nonverbal Accuracy (DANVA2; Nowicki, 2005); na test składają się dwie części: w *DANVA dzieci* prezentowane są próbki głosu dzieci, w *DANVA dorośli* – dorosłych;
- Skalę Inteligencji Emocjonalnej – Twarze (SIE-t; Matczak, Piekarska i Studniarek, 2005), służącą sprawdzaniu zdolności do identyfikacji emocjonalnej ekspresji mimicznej za pomocą etykiet słownych;
- Popularny Kwestionariusz Inteligencji Emocjonalnej (PKIE; Jaworowska i Matczak, 2005), obejmujący cztery podskale: akceptacji emocji, empatii, kontrolowania, rozumienia własnych emocji;
- trzy testy z Baterii Testów APIS-Z do pomiaru różnych aspektów inteligencji poznawczej: *zachowania* do pomiaru rozumienia reguł społecznych, *synonimy* badający zdolności werbalne, *przekształcenia liczb* do pomiaru zdolności abstrakcyjno-logicznych (Matczak, Jaworowska, Ciechanowicz i Stańczak, 2006);
- Skalę Wrażliwości Empatycznej (SWE; Kaźmierczak, Plopa i Retowski, 2007), opartą na kwestionariuszu empatii Davisa;
- Inwentarz Osobowości NEO-FFI (Zawadzki, Strelau, Szczepaniak i Śliwińska, 1998);
- Skalę Ustosunkowań Interpersonalnych (SUI; Stanik, 1994, 1998), w której poszczególne postawy interpersonalne, takie jak opiekuńcza, agresywna, narcystyczna, traktowane są jako rysy osobowości, na mocy teorii Leary'ego.

Prezentacja wyników

Wyniki korelacji TWES z miarami testowymi rozpoznawania i rozumienia emocji oraz inteligencji zawiera tabela 7, natomiast z kwestionariuszami samooceny inteligencji emocjonalnej i empatii tabela 8.

Choć autorki Testu Rozumienia Emocji nie zalecają używania wyników poszczególnych zadań (Matczak i Piekarska, 2011), jednak ze względu na specyfikę prowadzonych analiz, sprawdzona została korelacja nie tylko ogólnej

Tabela 7. Współczynniki korelacji Spearmana między TWES a testami umiejętności emocjonalnych i poznawczych. Wiek badanych w badaniu z wykorzystaniem SIE-t: 15 lat, w pozostałych grupach 16–18 lat

Test	N	TWES _t	TWES _{tP}	TWES _k
TRE całość	123	0,14	0,15	0,15
TRE: emocje w sytuacjach		0,31*** [0,07–0,52]	0,31*** [0,07–0,52]	-0,03
APIS zachowania	65	-0,14	-0,12	0,13
APIS synonimy	93	-0,05	-0,05	0,21# [-0,08–0,46]
APIS przekształcenia liczb	59	-0,14	-0,10	0,31# [-0,05–0,60]
DANVA dzieci	102	0,11	0,14	0,07
DANVA dorośli		0,16	0,17# [-0,10–0,42]	-0,02
DANVA całość		0,16	0,19# [-0,08–0,44]	0,04
SIE-t ^a	64	0,17	0,21# [-0,14–0,51]	0,00

TRE – Test Rozumienia Emocji; APIS – bateria testów inteligencji; DANVA – Diagnostic Analysis of Nonverbal Accuracy z podskalami ekspresji dzieci i dorosłych; SIE-t – Skala Inteligencji Emocjonalnej – Twarze.

W nawiasach kwadratowych podano granice przedziałów ufności dla współczynnika korelacji (na poziomie ufności 95%, obliczenia za pomocą kalkulatora ze strony: <http://karasiewicz.az.pl>).

^a Dane z: (Madera, 2009).

*** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; # $p < 0,1$.

Tabela 8. Wskaźniki korelacji Spearmana między TWES a samooceną inteligencji emocjonalnej i empatii

Kwestionariusz	TWES _t	TWES _{tP}	TWES _k
PKIE całość	0,19#	0,21#	-0,01
PKIE akceptacja	0,13	0,17	0,12
PKIE empatia	0,19	0,20#	-0,07
PKIE kontrola	0,16	0,15	-0,04
PKIE rozumienie	0,09	0,09	0,03
SWE całość	-0,23*** a -0,31** a	-0,22*** a -0,30** a	0,26*** a 0,28**
SWE: empatyczna troska	-0,17** -0,32** a	-0,15* -0,30** a	0,21*** a 0,25*
SWE: osobista przykrość	-0,27*** a -0,36*** a	-0,26*** a -0,35*** a	0,18** a 0,25*
SWE: przyjmowanie perspektywy	-0,03 -0,10	-0,02 -0,09	0,17** 0,10

PKIE – Popularny Kwestionariusz Inteligencji Emocjonalnej (N=79; wiek: 13–77; $M_{\text{wiek}}=35,5$); SWE – Skala Wrażliwości Empatycznej, badanie przeprowadzone w dwóch grupach: na górze – dane pozyskane dzięki uprzejmości p. Marcina Moronia z Instytutu Psychologii Uniwersytetu Śląskiego (N=256; wiek 17–18), na dole – dane z: Włodarczak, 2010 (N=90; wiek: 19–82; $M_{\text{wiek}}=49,1$; $SD_{\text{wiek}}=18,6$).

^a Współczynniki korelacji dla których lewa granica przedziału ufności (na poziomie ufności 95%) jest większa niż 0.

*** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; # $p < 0,1$.

punktacji TRE, ale także skali, w której badani wskazują najbardziej prawdopodobne emocje w opisanych sytuacjach. Ten właśnie wynik koreluje z TWES, obliczanym według klucza targetowego, mimo braku związku z całościową punktacją TRE.

Nie odnotowano istotnych statystycznie korelacji TWES z testami rozpoznawania emocji, jakimi są DANVA i SIE-t, ani z miarami umiejętności poznawczych. W odniesieniu do kwestionariusza empatii SWE pojawiły się związki dodatnie przy zastosowaniu klucza konsensualnego i ujemne przy kluczu targetowym.

Nie odnotowano istotnych korelacji między wynikami TWES a wymiarami osobowości w modelu Wielkiej Piątki ($N=115$; $M_{\text{wiek}}=39,9$, $SD_{\text{wiek}}=14,5$). W analizie związków z wymiarami Skali Ustosunkowań Interpersonalnych ($N=146$; $M_{\text{wiek}}=29,8$; $SD_{\text{wiek}}=12,7$) uzyskano wskaźniki korelacji na poziomie podobnym jak w NEO-FFI z wyjątkiem stylu współpracująco-przyjacielskiego, korelującego z TWES na poziomie $r_s=-0,17$ ($p<0,05$). Tabela z wynikami korelacji z wymiarami osobowości znajduje się w materiałach dodatkowych.

Dyskusja wyników

Uzyskane wyniki w dużym stopniu potwierdzają trafność TWES przy zastosowaniu klucza targetowego. Wprawdzie TWES nie koreluje z Testem Rozumienia Emocji, jednak przy kluczu targetowym związek istotny pojawił się z podtestem, badającym wnioskowanie o emocjach w opisanych sytuacjach, czyli z częścią sprawdzającą tę samą umiejętność co TWES. Brak istotnych korelacji z ogólnym wynikiem TRE, a także z SIE-t i DANVA wskazuje, że TWES mierzy wąsko rozumianą zdolność do wnioskowania o funkcjonowaniu emocjonalnym. Korelacja TWES_T z samooceną inteligencji emocjonalnej nie osiągnęła poziomu istotności, jednak ma wartość analogiczną do uzyskanej testem TEMINT w badaniu Choi i Kluemper (2012).

Stosowane dotychczas testy rozumienia emocji korelują z miarami inteligencji, przede wszystkim werbalnej (Matczak i Piekarska, 2011; Mayer, Salovey i Caruso, 2004; Śmieja i in., 2007), jednak przy wykorzystaniu TEMINT, czyli testu o konstrukcji podobnej do TWES, takich związków nie odnotowano (Blickle i in., 2009). W przypadku TWES nie ma istotnych korelacji z rozumieniem sytuacji społecznych, zdolnościami werbalnymi, ani zdolnościami abstrakcyjno-logicznymi. Wynik taki jest niezgodny z założeniem, że wnioskowanie o procesach emocjonalnych jest częścią inteligencji emocjonalnej, a ta z kolei wchodzi w zakres inteligencji ogólnej (por. Mayer i in., 1999; Roberts, Zeidner i Matthews, 2001). Być może odnotowane w innych badaniach korelacje są efektem tego, że MSCE-IT, TRE i TIE w większym stopniu wykorzystują elementy rozumowania ściśle poznawczego (nieemocjonalnego). Na podstawie dotychczasowych badań trudno tę kwestię rozstrzygnąć.

Zaskakujące mogą się wydawać negatywne związki TWES targetowego z wymiarami empatii, jednak podobny efekt odnotowany został w odniesieniu do TEMINT (Blickle i in., 2009). Zjawisko to, podobnie jak korelacja ujemna z postawą przyjacielsko-współpracującą w SUI, może być efektem współwystępowania dwóch elementów. Badani, którzy uważają, że świat relacji interpersonalnych pełen jest emocji, zaznaczają w TWES odpowiedzi pozytywne, zyskując wyższą punktację w kluczu konsensualnym, ale niską w targetowym. Jednocześnie, przeszacowując prawdopodobieństwo reakcji emocjonalnych, są przekonani o swojej wrażliwości i prospołecznym nastawieniu, zaznaczając więcej punktów w samoopisie empatii i przyjacielskiej postawy.

Wyniki TWES nie mają związku z wymiarami osobowości, co jest zgodne z założeniami Mayera i współpracowników (Mayer i in., 2004; Mayer, Roberts i Barsade, 2008) odnośnie umiejętności emocjonalnych wchodzących w zakres inteligencji emocjonalnej.

PODSUMOWANIE

Na podstawie prezentowanych powyżej wyników badań stwierdzić można, że Test Rozumienia Emocji w Sytuacji jest narzędziem o odpowiedniej dla badań grupowych rzetelności i trafności.

Wypełnienie arkusza zabiera osobom dorosłym około 10 minut i polega na zaznaczeniu w skali trzypunktowej, czy każda z wymienionych czterech emocji może pojawić się w opisanej sytuacji. Do obliczania wyników wykorzystać można dwa klucze: targetowy i konsensualny. W obu przypadkach punktacja oparta została na wynikach analiz związków sytuacja-emocja, jednak w przygotowaniu klucza targetowego badani przewidywali własne emocje w arkuszu pre-TWES, a w kluczu konsensualnym odnosili się do przypuszczalnych reakcji afektywnych innych osób, wypełniając arkusz TWES. Można przyjąć, że klucz targetowy w większym stopniu związany jest ze zdolnością do przewidywania rzeczywistych emocji, z uwzględnieniem tych reakcji, które raczej nie pojawią się w danej sytuacji, natomiast TWESk jest wskaźnikiem zgodności indywidualnych skryptów emocjonalnych z obrazem funkcjonowania afektywnego człowieka, obowiązującym w społeczeństwie. Wyniki uzyskane przy zastosowaniu obu kluczy korelują ze sobą ujemnie, ponieważ w TWESk dużą część punktów uzyskuje się dzięki wskazaniom emocji, które raczej nie pojawiają się w opisanych sytuacjach, natomiast wyższe wyniki w TWESk wiążą się z częstszym określaniem emocji jako prawdopodobnych. W interpretacji należy uwzględnić, że wyniki testu mogą być podatne na zniekształcenia wynikające z tendencji do zgadzania się, przy czym w TWESk tendencja ta obniża punktację, a w TWESk podwyższa. Ze względu na uzyskane współczynniki korelacji z miarami rozumienia emocji, wskazujące na trafność narzędzia, rekomendowane jest używanie klucza targetowego.

Współczynniki rzetelności α -Cronbacha są zbliżone dla obu kluczy. Na podstawie wartości α -Cronbacha przyjąć można, że TWES pozwala

na rzetelny pomiar zarówno u dorosłych jak i dzieci powyżej jedenastego roku życia, przy czym w odniesieniu do dziewcząt w wieku gimnazjalnym należy ostrożnie wykorzystywać wynik TWESk. Test może być wykorzystany zarówno do badania zmian rozwojowych, jak i w analizie związków rozumienia emocji z funkcjonowaniem poznawczym, zdrowotnym, interpersonalnym, z sukcesem edukacyjnym i zawodowym (por. np. Brackett, Rivers i Salovey, 2011; Mayer, Salovey i Caruso, 2004).

Brak związku z miarami inteligencji poznawczej wskazuje, iż umiejętność wnioskowania o emocjach nie jest jedynie zastosowaniem rozumowania logicznego lub wiedzy do specyficznego materiału, jakim jest sfera funkcjonowania emocjonalnego. Brak istotnych korelacji z wymiarami osobowości sugeruje, że zróżnicowania indywidualnego w zakresie wnioskowania o emocjach nie należy wiązać bezpośrednio z opisywanymi w ramach teorii osobowości, charakterystycznymi dla danej osoby formami reagowania.

Materiały dodatkowe można pobrać z:

<https://drive.google.com/file/d/0B8S631-LZmN6MH-d4RVV1bkNVMWs/view?usp=sharing>

Arkusz badawczy dla dorosłych:

<https://drive.google.com/file/d/0B8S631-LZmN6ZGs5X1ZFcy1sMGc/view?usp=sharing>

Arkusz dla badanych w wieku szkolnym:

<https://drive.google.com/file/d/0B8S631-LZmN6M3llSFIVbGNpSIU/view?usp=sharing>

W sprawie przeliczenia wyników proszę kontaktować się z autorem.

Badania prezentowane w artykule częściowo finansowane były w ramach projektu Instytutu Psychologii Uniwersytetu Śląskiego 06/BW/IPS/08

LITERATURA CYTOWANA

- Austin, E. J. (2009). A reaction time study of responses to trait and ability emotional intelligence test items. *Personality and Individual Differences*, 46(3), 381–383.
- Austin, E. J. (2010). Measurement of ability emotional intelligence: Results for two new tests. *British Journal of Psychology*, 101(3), 563–578.
- Austin, E. J., Farrelly, D., Black, C. i Moore, H. (2007). Emotional intelligence, Machiavellianism and emotional manipulation: Does EI have a dark side? *Personality and Individual Differences*, 43(1), 179–189.
- Barchard, K. A. i Russell, J. A. (2006). Bias in consensus scoring, with examples from ability emotional intelligence tests. *Psicothema*, 18 (Suplement), 49–54.
- Bar-On, R. M. (2006). The Bar-On model of emotional-social intelligence (ESI). *Psicothema*, 18(1), 13–25.
- Baron-Cohen, S. (2001). *Theory of mind in normal development and autism*. Prisme. 34, 174–183. Pobrano wrzesień, 2014, z: www.autism-community.com
- Berrocal, P., & Extremera, N. (2008). A review of trait meta-mood research. *Advances in Psychology Research*, 55, 17–45.
- Blickle, G., Momm, T., Liu, Y., Witzki, A. i Steinmayr, R. (2011). Construct validation of the Test of Emotional Intelligence (TEMINT): A two-study investigation. *European Journal of Psychological Assessment*, 27(4), 282–289.
- Blickle, G., Momm, T. S., Kramer, J., Mierke, J., Liu, Y. i Ferris, G. R. (2009). Construct and criterion-related validation of a measure of emotional reasoning skills: A two-study investigation. *International Journal of Selection and Assessment*, 17(1), 101–118.
- Brackett, M. A., Rivers, S. E. i Salovey, P. (2011). Emotional intelligence: implications for personal, social, academic, and workplace success. *Social and Personality Psychology Compass*, 5, 88–103.
- Brackett, M. A., Rivers, S. E., Shiffman, S., Lerner, N. i Salovey, P. (2006). Relating emotional abilities to social functioning: A comparison of self-report and performance measures of emotional intelligence. *Journal of Personality and Social Psychology*, 91(4), 780–795.
- Brackett, M. A. i Salovey, P. (2006). Measuring emotional intelligence with the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT). *Psicothema*, 18 (Suplement), 34–41.
- Brzeziński, J. (1999). *Metodologia badań psychologicznych*. Warszawa: Wydawnictwo Naukowe PWN.
- Burkitt, E. i Watling, D. (2015). How do children who understand mixed emotion represent them in freehand drawings of themselves and others? *Educational Psychology*, 1–21.
- Burnett, S., Thompson, S., Bird, G. i Blakemore, S.-J. (2011). Pubertal development of the understanding of social emotions: Implications for education. *Learning and Individual Differences*, 21(6), 681–689.
- Choi, S. i Kluemper, D. H. (2012). The relative utility of differing measures of emotional intelligence: Other-report EI as a predictor of social functioning. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology*, 62(3), 121–127.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, (10), 85–101.
- Davis, S. K. i Humphrey, N. (2014). Ability versus trait emotional intelligence: Dual influences on adolescent psychological adaptation. *Journal of Individual Differences*, 35(1), 54–62.
- Day, A. L. i Carroll, S. A. (2004). Using an ability-based measure of emotional intelligence to predict individual performance, group performance, and group citizenship behaviours. *Personality and Individual Differences*, 36(6), 1443–1458.
- Fehr, B. i Russell, J. A. (1984). Concept of emotion viewed from a prototype perspective. *Journal of Experimental Psychology: General*, 113, 464–486.
- Fine, S. E., Izard, C. E. i Trentacosta, C. J. (2006). Emotion situation knowledge in elementary school: Models of longitudinal growth and preschool correlates. *Social development*, 15(4), 730–751.
- Gignac, G. E. (2010). *Genos Emotional Intelligence Inventory technical manual (2nd edition)*. Sydney: NSW: Genos Press. Pobrano styczeń, 2014 z: <https://www.genosinternational.com/emotional-intelligence/technical-manual>
- Grubb III, W. L. i McDaniel, M. A. (2007). The fakability of Bar-On's Emotional Quotient Inventory Short Form: catch me if you can. *Human Performance*, 20(1), 43–59.
- Hansen, K., Lloyd, J. i Stough, C. (2009). Emotional intelligence and clinical disorders. *Assessing emotional intelligence: Theory, research, and applications* (ss. 219–237). New York: Springer Science, Business Media.

- Harter, S. i Buddin, B. J. (1987). Children's understanding of the simultaneity of two emotions: A five-stage developmental acquisition sequence. *Developmental Psychology*, 23(3), 388.
- Hoerger, M., Chapman, B. P., Epstein, R. M. i Duberstein, P. R. (2012). Emotional intelligence: A theoretical framework for individual differences in affective forecasting. *Emotion*, 12(4), 716–725.
- Izard, C. E. (2001). Emotional intelligence or adaptive emotions? *Emotion*, 1(3), 249–257.
- Janowski, M. (2015). Międzypokoleniowa nieadekwatność obrazu reakcji emocjonalnej jako potencjalne źródło nieporozumień. *Spoleczeństwo i Edukacja*, 16(1), 99–105.
- Jaworowska, A. i Matczak, A. (2005). *PKIE – Popularny Kwestionariusz Inteligencji Emocjonalnej*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Każmierczak, M., Plopa, M. i Retowski, S. (2007). Skala wrażliwości empatycznej. *Przegląd Psychologiczny*, 50(1), 9–24.
- Libbrecht, N. i Lievens, F. (2012). Validity evidence for the situational judgment test paradigm in emotional intelligence measurement. *International Journal of Psychology*, 47(6), 438–447.
- Lizeretti, N. P. i Extremera, N. (2011). Emotional intelligence and clinical symptoms in outpatients with Generalized Anxiety Disorder (GAD). *Psychiatric Quarterly*, 82(3), 253–260.
- MacCann, C. (2010). Further examination of emotional intelligence as a standard intelligence: A latent variable analysis of fluid intelligence, crystallized intelligence, and emotional intelligence. *Personality and Individual Differences*, 49(5), 490–496.
- MacCann, C. E. (2006). New approaches to measuring emotional intelligence. University of Sydney. School of Psychology. Pobrano wrzesień, 2009, z: <http://ses.library.usyd.edu.au/handle/2123/934>
- MacCann, C., Matthews, G., Zeidner, M. i Roberts, R. (2003). Psychological assessment of emotional intelligence: a review of self-report and performance-based testing. *International Journal of Organizational Analysis*, 3(11), 247–274.
- MacCann, C., Pearce, N. i Roberts, R. D. (2011). Emotional intelligence as assessed by situational judgment and emotion recognition tests: Building the nomological net. *Psychological Topics*, 20(3), 393–412.
- MacCann, C. i Roberts, R. D. (2008). New paradigms for assessing emotional intelligence: Theory and data. *Emotion*, 8(4), 540–551.
- Madera, K. (2009). *Wpływ inteligencji emocjonalnej na poczucie samotności, poczucie wsparcia rodzicielskiego oraz koleżeńskigo u młodzieży w okresie dorastania*. Niepublikowana praca magisterska. Uniwersytet Śląski, Katowice.
- Magnusson, D. (1991). *Wprowadzenie do teorii testów*. Warszawa: Państwowe Wydawnictwo Naukowe.
- Matczak, A., Jaworowska, A., Ciechanowicz, A. i Stańczak, J. (2006). *Bateria testów APIS-Z: podręcznik*. Warszawa: Pracownia Testów Psychologicznych Polskiego Towarzystwa Psychologicznego.
- Matczak, A. i Piekarska, J. (2011). *Test Rozumienia Emocji TRE. Podręcznik*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Matczak, A., Piekarska, J. i Studniarek, E. (2005). *Skala Inteligencji Emocjonalnej – Twarze (SIE-T): podręcznik*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Matthews, G., Zeidner, M. i Roberts, R. D. (2004). *Emotional intelligence: science and myth*. Cambridge: MIT Press.
- Maul, A. (2011). The factor structure and cross-test convergence of the Mayer–Salovey–Caruso model of emotional intelligence. *Personality and Individual Differences*, 50(4), 457–463.
- Mayer, J. D. (2000). Emotion, intelligence, and emotional intelligence. W: J. P. Forgas (Red.), *The handbook of affect and social cognition* (ss. 410–431). Mahwah, NJ: Lawrence Erlbaum & Associates.
- Mayer, J. D., Caruso, D. R. i Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, 4(27), 267–298.
- Mayer, J. D., Roberts, R. D. i Barsade, S. G. (2008). Human abilities: Emotional intelligence. *Annual Review of Psychology*, 59(1), 507–536.
- Mayer, J. D. i Salovey, P. (1997). What is emotional intelligence? W: P. Salovey i D. Sluyter (Red.), *Emotional development and emotional intelligence: Implications for educators* (ss. 3–31). New York: Basic Books.
- Mayer, J. D., Salovey, P. i Caruso, D. R. (2000). Models of emotional intelligence. *The handbook of intelligence* (ss. 396–420). New York: Cambridge University Press.

- Mayer, J. D., Salovey, P. i Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, 15(1), 197–215.
- Mayer, J. D., Salovey, P., Caruso, D. R. i Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion*, 3(1), 97–105.
- Nowicki, S., Jr. (2005). Manual for the receptive tests of the Diagnostic Analysis of Nonverbal Accuracy 2 (DANVA2). Unpublished manuscript. Atlanta, Emory University.
- O'Connor, R. M. i Little, I. S. (2003). Revisiting the predictive validity of emotional intelligence: self-report versus ability-based measures. *Personality and Individual Differences*, 35(8), 1893–1902.
- Petrides, K. V. i Furnham, A. (2001). Trait emotional intelligence: psychometric investigation with reference to established trait taxonomies. *European Journal of Personality*, 15(6), 425–448.
- Reissland, N. (1985). The development of concepts of simultaneity in children's understanding of emotions. *Journal of Child Psychology and Psychiatry*, 26(5), 811–824.
- Roberts, R. D., Schulze, R., O'Brien, K., MacCann, C., Reid, J. i Maul, A. (2006). Exploring the validity of the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) with established emotions measures. *Emotion*, 6(4), 663–669.
- Roberts, R. D., Zeidner, M. i Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion*, 1(3), 196–231.
- Rode, J. C., Mooney, C. H., Arthaud-day, M. L., Near, J. P., Rubin, R. S., Baldwin, T. T. i Bommer, W. H. (2008). An examination of the structural, discriminant, nomological, and incremental predictive validity of the MSCEIT V2.0. *Intelligence*, 36(4), 350–366.
- Roseman, I. J., Spindel, M. S. i Jose, P. E. (1990). Appraisals of emotion-eliciting events: Testing a theory of discrete emotions. *Journal of Personality and Social Psychology*, 59(5), 899.
- Saarni, C. (1999). *The Development of Emotional Competence*. New York, London: Guilford Press.
- Sagan, A. (2003). *Model pomiarowy satysfakcji i lojalności*. StatSoft Polska. Pobrano czerwiec, 2014 z: www.statsoft.pl/Portals/0/Downloads/pomiarowy.pdf
- Salguero, J. M., Extremera, N. i Fernández-Berrocal, P. (2012). Emotional intelligence and depression: The moderator role of gender. *Personality and Individual Differences*, 53(1), 29–32.
- Salovey, P., Mayer, J. D., Goldman, S. L., Turvey, C. i Palfai, T. P. (1995). Emotional attention, clarity, and repair: Exploring emotional intelligence using the Trait Meta-Mood Scale. *Emotion, disclosure i health* (ss. 125–154). Washington: American Psychological Association.
- Sharma, S., Gangopadhyay, M., Austin, E. i Mandal, M. K. (2013). Development and validation of a situational judgment test of emotional intelligence. *International Journal of Selection and Assessment*, 21(1), 57–73.
- Siegling, A. B., Saklofske, D. H., Vesely, A. K. i Nordstokke, D. W. (2012). Relations of emotional intelligence with gender-linked personality: Implications for a refinement of EI constructs. *Personality and Individual Differences*, 52(7), 776–781.
- Smith, A. (2006). Cognitive empathy and emotional empathy in human behavior and evolution. *The Psychological Record*, 56(1), 3–21.
- Stanik, J. M. (1994). *Skala Ustosunkowań Interpersonalnych (SUI)*. Kielce: Wydawnictwo Szumacher.
- Stanik, J. M. (1998). Funkcjonowanie osobowości transseksualistów typu K/M w świetle badań testem SUI. *Acta Universitatis Palackianae Olomucensis. Psychologica*, 32, 87–106.
- Śmieja, M., Orzechowski, J. i Beauvale, A. (2007). TIE – Test Inteligencji Emocjonalnej. *Studia Psychologiczne*, 54(1), 80–99.
- Śmieja, M., Orzechowski, J. i Stolarski, M. S. (2014). TIE: An ability test of emotional intelligence. *PLoS ONE*, 9(7), e103484.
- Tonks, J., Williams, W. H., Frampton, I., Yates, P. i Slater, A. (2007). Assessing emotion recognition in 9–15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. *Brain Injury*, 21(6), 623–629.
- Trentacosta, C. J., Izard, C. E., Mostow, A. J. i Fine, S. E. (2006). Children's emotional competence and attentional competence in early elementary school. *School Psychology Quarterly*, 21(2), 148.
- Whitman, D. S., Van Rooy, D. L., Viswesvaran, C. i Alonso, A. (2008). The susceptibility of a mixed model measure of emotional intelligence to faking: a Solomon four-group design. *Psychology Science*, 50(1), 44.
- Widen, S. C. i Russell, J. A. (2011). In building a script for an emotion, do preschoolers add its cause before

- its behavior consequence? *Social Development*, 20, 471–485.
- Williams, C., Daley, D., Burnside, E. i Hammond-Rowley, S. (2009). Measuring emotional intelligence in preadolescence. *Personality and Individual Differences*, 47(4), 316–320.
- Włodarczak, E. (2010). *Poziom empatii a wiek, płeć, posiadanie dzieci i aktywność zawodowa*. Niepublikowana praca magisterska. Uniwersytet Śląski, Katowice.
- Zawadzki, B., Strelau, J., Szczepaniak, P. i Śliwińska, M. (1998). *Inwentarz Osobowości NEO-FFI Costy i McCrae. Adaptacja polska*. Warszawa: Pracownia Testów Psychologicznych PTP.
- Zeidner, M., Shanizinovich, I., Matthews, G. i Roberts, R. (2005). Assessing emotional intelligence in gifted and non-gifted high school students: Outcomes depend on the measure. *Intelligence*, 33(4), 369–391.

Maciej Janowski

University of Silesia

THE DEVELOPMENT AND PSYCHOMETRIC PROPERTIES OF THE TEST OF EMOTION INFERENCE FROM SITUATIONS (TWES)

ABSTRACT

The ability to infer others' emotions from situational cues has been studied in psychology in the context of emotional intelligence, empathy and theory of mind. The article presents the development, reliability and validity of a new measure of this skill. The Test of Emotion Inference from Situations (TWES) consists of 14 short stories and 4 emotions for every episode. Subjects were asked to assess the probability of emotional reactions in given situations. Consensus scores and target scores may be used for the TWES – in both cases the test was internally consistent and reliable (RMSEA < .05; Cronbach's α > .7). Correlational studies using emotional ability tests, tests of cognitive intelligence, emotional intelligence self-assessment questionnaires and personality inventories confirmed validity of the target scoring. The test may be used in scientific research with people over the age of eleven.

Keywords: emotional skills, understanding of emotion, emotional intelligence