

dr hab. Roman Dolata

Wydział Pedagogiczny UW
Instytut Badań Edukacyjnych
Zespół badawczy EWD

Artur Pokropek

Instytut Filozofii i Socjologii PAN
Zespół badawczy EWD

Motywacja a wynik testu z nauk przyrodniczych Studium na przykładzie PISA 2006¹

Problem

Motywacja wpływa na wyniki testów osiągnięć szkolnych na dwa główne sposoby. Po pierwsze, jako względnie stały mechanizm regulacyjny wpływa na procesy uczenia się szkolnego. Motywacja tego typu może mieć ogólny, niezależny od przedmiotu, charakter lub poprzez mechanizm zainteresowania wiązać się z daną tylko dziedziną nauki szkolnej. W poniższym tekście zajmiemy się motywacją do nauki przedmiotów przyrodniczych. Po drugie, motywacja jako zmienny mechanizm zależny od czynników sytuacyjnych wpływa bezpośrednio na wykonanie zadań testowych. Ten typ motywacji nazywać będziemy motywacją testową (*test-taking motivation*). Motywacja testowa determinuje podjęcie zadania, wysiłek poznawczy i wytrwałość w dążeniu do rozwiązania zadań testowych.

W tym artykule zajmiemy się problemem motywacji w testach o małym znaczeniu dla rozwiązującego (*low-stake test*) na przykładzie testu PISA z zakresu nauk przyrodniczych (*science*). Podjęte będą dwa problemy. Po pierwsze, na ile wyniki w testach PISA (*science*) są zdeterminowane czynnikami motywacyjnymi związanymi z uczeniem się przedmiotów przyrodniczych i czy model zależności *motywacja – wyniki testu* jest taki sam we wszystkich krajach. Po drugie, czy wyniki testu PISA są zdeterminowane motywacją testową i – analogicznie – czy zależność ta jest kulturowo niezmienna. Z drugim problemem wiąże się podjęte w referacie zagadnienie metodologiczne – jak mierzyć motywację testową? Zwykle motywację testową mierzy się na podstawie danych samoopisowych pochodzących od osób testowanych. W referacie przedstawimy wyniki zastosowania nowej metody wskaźnikowania motywacji testowej na podstawie rekordu wykonania testu. Metoda ta bazować będzie na informacji o niepodjęciu przez rozwiązującego zadań testowych.

¹ Tekst powstał w ramach projektu badawczego N N114 446336 MNiSzW realizowanego przez Centrum Analiz Ekonomiczno-Społecznych CASE.

Główne podejścia teoretyczne do badania motywacji

Współczesne ujęcia motywacji wyrastają z teorii motywacji typu *Oczekiwanie x Wartość* (Expectancy x Value) oraz z koncepcji poczucia sprawstwa (*self-efficacy*). Zgodnie z podejściem *Oczekiwanie x Wartość* siła motywacji zależy od iloczynu oczekiwania sukcesu i jego wartości. Przykładem takiego podejścia do motywacji jest klasyczna koncepcja J.W. Atkinsona (1964). Badacz ten rozdziela czynnik *Wartość* na dwie składowe: motyw i przynętę. Dodatkowo Atkinson komplikuje obraz, twierdząc, że mogą równocześnie działać dwie tendencje – dążenie do sukcesu i dążenie do uniknięcia porażki. W wypadku dążenia do sukcesu podstawowe twierdzenie teorii Atkinsona można opisać za pomocą poniższej formuły.

$$T_s = M_s \times P_s \times I_s$$

We wzorze tym T_s oznacza siłę tendencji do sukcesu (odpowiednik pojęcia siły motywacji). Symbole M_s i I_s opisują wartość, przy czym pierwszy czynnik odnosi się do wewnętrznego aspektu wartości (motyw, czyli względnie stała dyspozycja jednostki do dążenia do sukcesu), natomiast drugi czynnik dotyczy zewnętrznego aspektu wartości (przynęta, czyli cechy obiektu spełniającego). Symbol P_s dotyczy oczekiwania odniesienia sukcesu, które Atkinson ujmuje jako prawdopodobieństwo. Siła motywacji do odniesienia sukcesu jest zatem iloczynem siły uogólnionego motywu sukcesu, spostrzeganej ważności danego zadania oraz prawdopodobieństwa sukcesu. O ile siła uogólnionego motywu dążenia do sukcesu jest względnie stałą cechą jednostki, to dwa pozostałe czynniki zmieniają się pod wpływem czynników sytuacyjnych.

Analogiczna formuła opisuje motywację uniknięcia porażki, którą Atkinson oznacza symbolem T_{\cdot} . Siła motywacji wypadkowej zależy (w pewnym uproszczeniu) od wypadkowej T_s i T_{\cdot} .

Współcześnie podejście *Oczekiwanie x Wartość* do problemu motywacji do uczenia się typu szkolnego odnajdujemy w pracach A. Wigfielda i J. Ecclessa (2000). Zgodnie z ich poglądami **oczekiwanie** przez ucznia sukcesu w danym zadaniu zależy od przekonania o posiadaniu kompetencji i spostrzeganej trudności zadania. Z kolei **wartość** zależy od czterech czynników: wartości zadania z perspektywy osiągnięć (*attainment value*, np. na ile ważne jest dobre wykonanie tego zadania), wartości wewnętrznej zadania (*intrinsic value*, np. zainteresowanie zadaniem, przyjemność z jego wykonywania), wartości instrumentalnej (*utility value*, np. na ile to zadanie wpasowuje się w plany na przyszłość) oraz od spostrzeganych kosztów (np. co można by robić innego, gdyby porzucić zadanie). W stosunku do teorii Atkinsona rozbudowaniu uległa konceptualizacja zarówno czynnika oczekiwania, jak i wartości. Oczekiwanie jest funkcją przekonania o posiadaniu kompetencji oraz spostrzeganej trudności zadania. (Co ciekawe, w obiektywistycznej wersji model ten odnajdujemy w podejściu IRT do skalowania testów.) Z kolei czynnik wartości został rozpisany na 4 elementy. Zanikła natomiast dialektyka dążenia do sukcesu i dążenia do uniknięcia porażki.

Element poczucia kompetencji znajduje pełniejsze ujęcie w teorii poczucia sprawstwa (Bandura, 1982). Poczucie sprawstwa to subiektywna ocena tego, czy posiada się odpowiednie umiejętności do wykonania zadań określonego

typu. Konstrukty ten w odróżnieniu od samooceny jest silniej powiązany z typem zadań i w związku z tym ma mniej zgeneralizowany charakter. Jeżeli w danej dziedzinie spostrzegam swoje umiejętności jako niskie, to niezależnie od obiektywnego poziomu kompetencji gorzej radzę sobie z zadaniami. Osoby o niskiej *self-efficacy* pracują mniej wytrwale, mniej wysiłku wkładają w rozwiązanie zadania, stosują bardziej prymitywne strategie radzenia sobie z zadaniem (Dembo, 1994).

Powyższe ujęcia teoretyczne znajdują zastosowanie zarówno w analizach motywacji do uczenia się szkolnego, jak i motywacji testowej. Podsumowując rozważania teoretyczne, można zaproponować następującą listę czynników motywacyjnych, swoistych dla nauki szkolnej i motywacji testowej w zakresie przedmiotów przyrodniczych:

- przypisywana przez ucznia naukom przyrodniczym wartość społeczna i osobista;
- zainteresowanie tym, czego dotyczy uczenie się przedmiotów przyrodniczych, i przyjemność towarzysząca uczeniu się;
- użyteczność przedmiotów przyrodniczych dla dalszej edukacji i przeszłej kariery zawodowej;
- względnie trwałe przekonanie o tym, na ile jestem kompetentny w dziedzinie nauk przyrodniczych.

Powyższa lista pokrywa się z założeniami pomiarowymi motywacji w badaniu PISA/OECD 2006.

Motywacja testowa w testach o niewielkim znaczeniu dla rozwiązującego

We współczesnych systemach edukacyjnych coraz większego znaczenia nabierają wyniki testów służących ewaluacji instytucji edukacyjnych i polityce oświatowej. Zwykle do tego celu używa się specjalnych testów, a ich wyniki nie wpływają na szkolne losy uczniów. Czasami rozwiązujący nie poznają nawet swojego wyniku. Dobrym przykładem takich testów są narzędzia pomiarowe stosowane w amerykańskim programie NAEP czy OECD PISA. Używanie tego typu testów, zwanych w literaturze przedmiotu *low-stake test*, wiąże się z dwoma istotnymi problemami. Po pierwsze, wyniki takich testów prawdopodobnie zaniżają faktyczny poziom umiejętności uczniów, a tym samym prowadzą do niedoszacowania efektywności edukacji (Wise, DeMars, 2005). Po drugie, rodzą się wątpliwości co do teoretycznej trafności tych testów. Jeżeli testy *low-stake* i *high-stake* mają inny układ empirycznych korelatów, to być może (trochę) co innego mierzą (Sundre, Wise, 2003; Wise, DeMars, 2005). Może to mieć rozliczne konsekwencje. Przede wszystkim może prowadzić do nietrafnych ocen ewaluacyjnych poszczególnych instytucji czy całych systemów edukacyjnych. Dlatego tak ważne jest dobre zrozumienie specyfiki tego typu testów. Jednym z kluczowych dla zrozumienia specyfiki testów *low-stake* fenomenów jest motywacja, zarówno ogólna motywacja szkolna związana z nauką danego/danych przedmiotów, jak i specyficzna motywacja testowa (*test taking motivation*) decydująca o wysiłku włożonym w rozwiązywanie zadań danego testu.

W odniesieniu do motywacji testowej w testach *low-stake* omawiane powyżej koncepcje teoretyczne motywacji prowadzą do wielu wniosków. Zwróćmy

uwagę na jeden z nich. Na pierwszy rzut oka konstatacja jest banalna – testy *low-stake*, mając mniejszą wartość przynętową, mogą wywoływać słabszą motywację. Z drugiej jednak strony, zgodnie z teorią Atkinsona, porażka w takich testach jest mniej zagrażająca, czyli tendencja do uniknięcia porażki powinna być słabsza. Ponieważ u większości osób dominuje tendencja do osiągnięcia sukcesów, więc pierwszy efekt jest silniejszy i generalnie wyniki w testach *low-stake* powinny być niższe niż w testach *high-stake*. Nie jest to jednak – prawdopodobnie – proste, addytywne obniżenie wyniku u wszystkich rozwiązujących. Ponieważ siła motywu sukcesu i motywu unikania porażki nie jest w prosty sposób skorelowana, spadek wykonania testu zależy od konfiguracji tych dwóch motywów. To ma konsekwencje dla trafności testów *low-stake*. Problem polega nie tylko na tym, że testy tego typu zaniżają oszacowanie umiejętności rozwiązujących, ale testy te mogą mierzyć co innego niż testy *high-stake*. Te rozważania są tylko ilustracją problemów z teoretyczną trafnością testów *low-stake*.

Badanie motywacji testowej

Badanie motywacji testowej jest wyzwaniem zarówno teoretycznym, jak i praktycznym – badacze tworzą testy motywacji testowej, by zrozumieć to zjawisko; wiele firm testowych tworzy takie narzędzia, by kontrolować trafność swoich pomiarów (Wolf and Smith, 1995; Wise i DeMars, 2005; Cole, Bergin, Whittaker, 2008). Generalnie można wyróżnić dwa podejścia: eksperymentalne i korelacyjne. W podejściu eksperymentalnym manipuluje się poziomem motywacji testowej, w korelacyjnym tworzy się testy motywacji testowej.

W podejściu eksperymentalnym można próbować manipulować motywacją testową poprzez wprowadzenie znaczących dla rozwiązujących konsekwencji, dostarczanie zachęt materialnych (*incentives*) za poprawne wykonanie zadań testowych, zmniejszanie pracochłonności zadań testowych, wykorzystywanie bardziej interesujących zadań, manipulowanie trudnością zadań tak, by ich spostrzegana trudność bliska była 50% szansie wykonania zadania i w końcu przez zapowiedź udzielenia informacji zwrotnej o poziomie wykonania testu (Baumert, Demmrich, 2001; DeMars, 2000). Skuteczność tych zabiegów zwykle jest niewielka lub zależna od cech testowanych.

W podejściu korelacyjnym stosuje się szereg testów motywacji. Wszystkie je łączą dwie cechy. Opierają się na samoopisie i są przeprowadzane po zakończeniu testowania. Zwykle mają też wielowymiarowy charakter. Na przykład *Student Opinion Survey* (Sundre, 1999) mierzy dwa aspekty motywacji testowej: *Reported Effort* i *Perceived Importance of Test*.

Nowe podejście do badania motywacji testowej

W naszej pracy przejęliśmy inne podejście do wskaźnikowania motywacji testowej. Pomiary oparte na samoopisie dokonywanym przez badanych po zakończeniu testowania są narażone na „zanieczyszczenie” wynikające z samooceny tego, jak rozwiązujący wypadł w teście. Na przykład, gdy respondent ma poczucie, że z testem sobie nie poradził, może próbować racjonalizować porażkę, obniżając ocenę wysiłku włożonego w wykonanie zadań.

Alternatywą wobec testów samoopisowych jest zbudowanie wskaźnika motywacji testowej na podstawie pełnego rekordu wykonania testu. Można próbować wykorzystywać informację o czasie wykonania testu, odsetku zadań podjętych, różnicy między wynikiem w zadaniach otwartych i zamkniętych lub różnicy między wynikiem uzyskanym w pierwszej i drugiej części testu (spadek motywacji). W naszych analizach wykorzystaliśmy drugą możliwość, czyli informację o podjęciu zadania testowego.

Dla każdego zadania i każdej osoby wykonującej test oszacowano prawdopodobieństwo podjęcia zadania przy kontroli jego trudności. Następnie prawdopodobieństwa te uśredniano dla każdego badanego i tak powstawał wskaźnik motywacji testowej. Różnica w stosunku do prostego wskaźnika odsetka podjętych przez badanego zadań polegała na tym, że kontrolowano trudność zadań.

Aby uzyskać opisywany wyżej efekt, użyty został dwupoziomowy model logistyczny z efektem losowym dla stałej, gdzie pierwszym poziomem analizy były zadania testowe (oznaczone indeksem t) zagnieżdżone w drugim poziomie modelowania – uczniach (oznaczone indeksem i).

Poziom 1 (zadanie):

$$\begin{aligned}\Pr(ODP_{it} = 1 | \beta) &= \varphi_{it} \\ \text{Log}[\varphi_{it} / (1 - \varphi_{it})] &= \eta_{it} \\ \eta_{it} &= \pi_{0i} + \pi_{1i}(TRU_{it}) + \pi_{2i}(ZAM_{it})\end{aligned}$$

Poziom 2 (uczeń):

$$\begin{aligned}\pi_{0i} &= \beta_{00} + r_{0i} \\ \pi_{1i} &= \beta_{10} \\ \pi_{2i} &= \beta_{20}\end{aligned}$$

Zmienną zależną w tym modelu jest zmienna zero-jedynkowa ODP_{it} przyjmująca wartość 0 dla t -tego zadania rozwiązywanego przez i -tego ucznia, gdy zadanie zostało opuszczone, a 1, gdy próba taka została podjęta niezależnie od tego, czy zadanie zostało rozwiązane poprawnie, czy nie. W modelu modelujemy prawdopodobieństwa warunkowe ze względu na zmienne niezależne, co znalazło odzwierciedlenie w pierwszym zapisie z równania z poziomu pierwszego definiującego warunkowe prawdopodobieństwo jako φ_{it} . Jako że posługujemy się modelem logitowym, warunkowe prawdopodobieństwo podczas modelowania wyrażane jest w wartościach *logodds*, czyli logarytmie naturalnym szans sukcesu η_{it} . Dzięki temu przekształceniu zmienna zależna zawiera się w przedziale od $+\infty$ do $-\infty$, co ułatwia modelowanie regresyjne. Model regresji pierwszego poziomu wyrażony jest w trzeciej liniije opisującej równanie poziom zadania. Można je odczytać następująco: zlogarytmizowane szanse podjęcia zadania zależą od stałej regresji π_{0i} , parametru związanego z trudnością pytania π_{1i} , gdzie TRU_{it} to zmienna charakteryzująca trudność pytania (w metryce IRT) oraz od wartości parametru π_{2i} przy zmiennej charakteryzującej to, czy pytanie jest zamknięte (1), czy otwarte (2). Prawdopodobieństwo (czy precyzyjniej zlogarytmizowane szanse) odpowiedzi na dane pytanie uzależnione

jest również od cechy ucznia r_{0j} , która różnicuje stałą regresji z pierwszego poziomu dla różnych uczniów. Jest to jedyny parametr z drugiego poziomu, w modelu założyliśmy bowiem, iż trudność zadania i jego typ związane są z prawdopodobieństwem podjęcia zadania dla każdego ucznia w taki sam sposób. Model ten estymowany był na pełnym zbiorze danych OECD PISA, przy czym użyto tutaj zrównoważonego ważenia, tak iż każdy kraj znaczył tyle samo podczas estymacji modelu niezależnie od liczby uczniów objętych badaniem. W trakcie szacowania modelu każdemu uczniowi wyliczona została wartość r_{0j} , którą dalej traktować będziemy jako wskaźnik motywacji testowej. Oszacowanie parametrów powyższego modelu dla danych z PISA-OECD 2006 przedstawione zostaną w rozdziale prezentującym wyniki analiz.

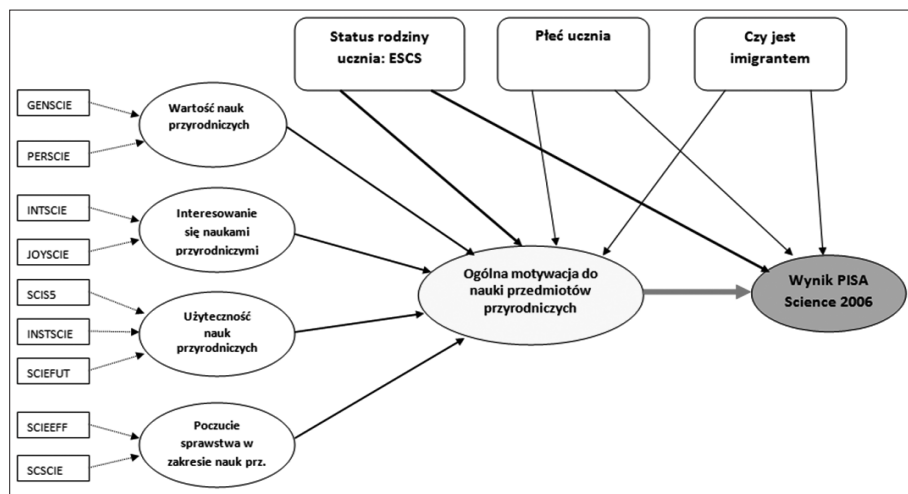
Motywacja do nauki przedmiotów przyrodniczych a wyniki testu PISA-OECD 2006, analiza porównawcza

W analizie związku między motywacją do nauki przedmiotów przyrodniczych a wynikami testu wykorzystano skalę z badania PISA-OECD 2006. Poniższa tabela zawiera krótki opis skal motywacji oraz listę zmiennych kontrolnych (pełny opis patrz: PISA 2006, Technical Report).

Tabela 1. Zmienne opisujące motywację do nauki przedmiotów przyrodniczych i zmienne kontrolne

Nazwa zmiennej	Skala (indeks) i przykładowa pozycja
GENSCIE	Ogólna wartość przypisywana naukom przyrodniczym <i><Broad science> is valuable to society</i>
PERSCIE	Osobista wartość przypisywana naukom przyrodniczym <i>I find that <broad science> helps me to understand the things around me</i>
INTSCIE	Interesowanie się nauką przedmiotów przyrodniczych <i>How much interest do you have in learning about the following <broad science> topics?</i>
JOYSCIE	Przyjemność towarzysząca nauce przedmiotów przyrodniczych <i>I generally have fun when I am learning <broad science> topics.</i>
SCIS5	Przewidywanie, że w wieku lat 30 będzie się wykonywało zawód powiązany z naukami przyrodniczymi
INSTSCIE	Instrumentalna motywacja do nauki przedmiotów przyrodniczych <i>Making an effort in my <school science> subject(s) is worth it because this will help me in the work I want to do later on</i>
SCIEFUT	Przewidywanie, że kariera zawodowa będzie się wiązała z naukami przyrodniczymi <i>I would like to work in a career involving <broad science></i>
SCIEEFF	Poczucie sprawstwa w zakresie nauk przyrodniczych <i>How easy do you think it would be for you to perform the following tasks on your own? -Describe the role of antibiotics in the treatment of disease</i>
SCSCIE	Samoocena powiązana z przedmiotami przyrodniczymi <i>I can usually give good answers to <test questions> on <school science> topics</i>
ESCS	Wskaźnik statusu ekonomiczno-społeczno-kulturowego rodziny ucznia
IMMIG	Bycie imigrantem
	Płeć

Związek między motywacją a wynikiem testu oszacowano za pomocą następującego modelu strukturalnego (SEM):



Schemat 1. Założony układ zależności w modelowaniu strukturalnym (SEM)

Współczynniki ścieżkowe oszacowano dla wszystkich krajów biorących udział w badaniu PISA-OECD 2006. Ze względu na brak miejsca zaprezentujemy tylko kluczowe wyniki uzyskane w analizach. Najważniejszy jest wskaźnik dla relacji *ogólna motywacja testowa* → *wynik testu PISA Science 2006*. Medialna wartość dla przeanalizowanych krajów wyniosła 0,26. Oznacza to umiarkowany, pozytywny związek. Interesująca jest jednak znaczna międzykrajowa zmienność tego wskaźnika. W Kirgistanie notujemy zależność negatywną (-0,20), w Islandii najwyższą pozytywną (0,49).

Analizując modele strukturalne dla zbadanych krajów, odnajdujemy dwie wyraźne prawidłowości:

(a) jeżeli w jakimś kraju obserwujemy silną zależność między motywacją a wynikiem testu, to równocześnie notujemy relatywnie wysoką korelację ESCS rodziny ucznia z motywacją oraz relatywnie niską zależność między ESCS rodziny ucznia a wynikami testu;

(b) im silniejszy w danym kraju związek między motywacją a wynikami testu, tym wyższy średni wynik testu.

Przyjrzyjmy się bliżej pierwszej prawidłowości. Tabela 2. przedstawia wyniki dla 16 krajów, w których związek między ogólną motywacją do nauki przedmiotów przyrodniczych a rezultatem w teście PISA-OECD Science 2006 jest relatywnie słaby. Tabela 3. zawiera natomiast wartości oszacowań współczynników ścieżkowych dla 10 krajów z relatywnie silną zależnością motywacja → wynik testu.

Tabela 2. Kraje, w których notujemy słaby związek między motywacją a wynikiem testowym

Kraje		Współczynniki ścieżkowe		
		motywacja → wynik testu	ESCS → wynik testu	ESCS → motywacja
południowo- amerykańskie	Kolumbia	-0,05	0,33	-0,06
	Argentyna	0,01	0,44	-0,04
	Brazylia	0,02	0,41	-0,01
	Urugwaj	0,05	0,44	0,01
	Meksyk	0,08	0,40	-0,03
	Chile	0,13	0,46	0,13
postkomunistyczne	Kirgistan	-0,20	0,28	-0,12
	Serbia	0,00	0,36	-0,05
	Azerbejdżan	0,03	0,21	0,03
	Rosja	0,07	0,28	0,02
	Rumunia	0,08	0,41	0,13
	Łotwa	0,13	0,30	0,10
	Węgry	0,16	0,44	0,14
	Słowacja	0,17	0,41	0,13
	Czechy	0,18	0,37	0,10
	Polska	0,18	0,36	0,13

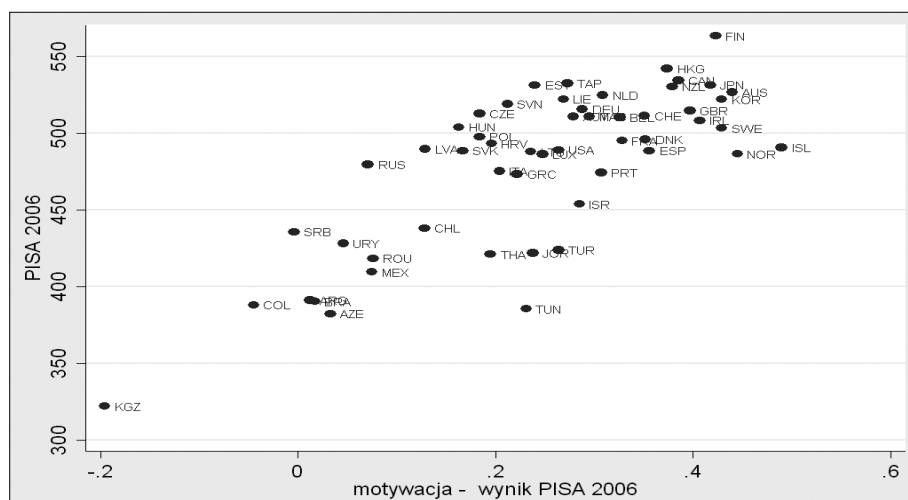
Tabela 3. Kraje, w których notujemy silny związek między motywacją a wynikiem testowym

Kraje		Współczynniki ścieżkowe		
		motywacja → wynik testu	ESCS → wynik testu	ESCS → motywacja
anglosaskie	Kanada	0,39	0,21	0,22
	Wielka Brytania	0,40	0,28	0,22
	Irlandia	0,41	0,24	0,28
	Australia	0,44	0,24	0,22
skandynawskie	Finlandia	0,42	0,17	0,24
	Szwecja	0,43	0,18	0,28
	Norwegia	0,45	0,15	0,25
	Islandia	0,49	0,12	0,26
azjaty- ckie	Japonia	0,42	0,19	0,19
	Korea Płd.	0,43	0,19	0,22

Analiza wyników pozwala wyróżnić dwa wzorce zależności. W pierwszym status rodziny ucznia (ESCS) silnie, bezpośrednio wpływa na wynik w teście. W drugim wzorcu obserwujemy silny związek ESCS z motywacją oraz silną relację między motywacją i wynikiem testu, natomiast bezpośredni wpływ ESCS na wyniki jest słabszy. Można powiedzieć, że w pierwszym wzorcu zależności status społeczny rodziny ucznia bezpośrednio odciska się na wynikach, w drugim modelu mediatorem jest motywacja do nauki przedmiotów przyrodniczych.

Niezwykle ciekawy jest silny związek tych dwóch wzorców z geopolityczną charakterystyką krajów. Model związku bezpośredniego między statusem rodziny ucznia a wynikiem testu występuje głównie w krajach południowoamerykańskich i postkomunistycznych. Drugi wzorec zależności obserwujemy głównie w krajach anglosaskich, skandynawskich i azjatyckich. To problem wymagający oddzielnych, rozbudowanych analiz.

Przejdźmy do drugiej prawidłowości, czyli im silniejszy w danym kraju związek między motywacją do nauki przedmiotów przyrodniczych a wynikami testu, tym wyższy średni wynik testu. Związek ten obrazuje poniższy wykres.



Wykres 1. Siła związku ogólna motywacja do nauki przedmiotów przyrodniczych → wynik testu PISA-OECD Science 2006 a średni wynik w tym teście. Jednostka analizy: kraj

Widzimy silny, liniowy związek między analizowanymi cechami krajów. Nawet po usunięciu wartości odstającej, Kirgistanu, korelacja wynosi $r=0,72$.

Motywacja testowa a wyniki testu PISA-OECD 2006

Oszacowanie parametrów wcześniej opisanego modelu do szacowania wskaźnika motywacji testowej zawiera poniższa tabela.

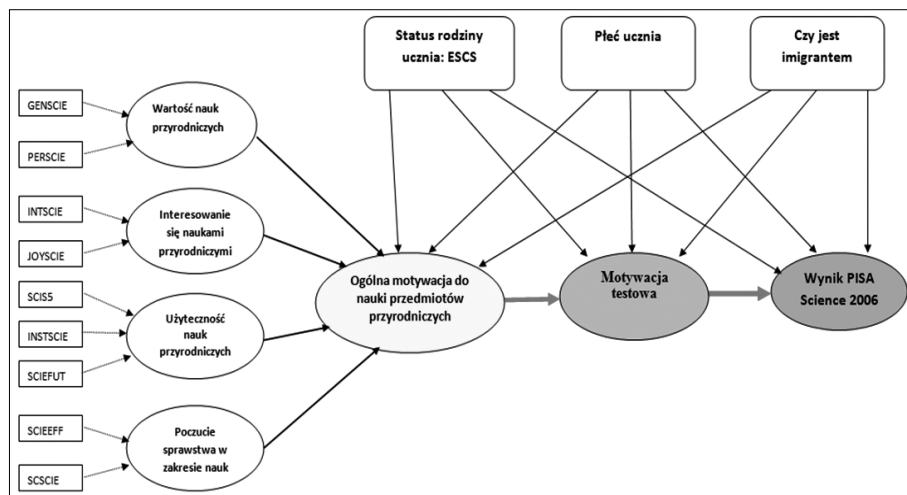
Tabela 4. Wyniki estymacji efektów stałych i efektu losowego (HLM6). Model średniej populacyjnej z odpornymi (robust) błędami standardowymi (tj. z poprawką na heteroscedastyczność oraz odstępstwa od normalności rozkładu reszt)

	współczynnik	błąd standardowy	statystyka t	poziom istotności	iloraz szans
stała	1,61	0,01	383,6	0,01	5,02
trudność zadania	-0.10	0,01	-58,8	0,01	0,90
zadanie zamknięte	1,19	0,01	214,2	0,01	3,30
	odchylenie standardowe	składnik wariancji	stopnie swobody	χ^2	poziom istotności
efekt losowy	1,71	2,92	398300	2740045	0,00

Wyniki dotyczące efektów stałych są dość oczywiste, ale potwierdzają sensowność modelu. Mianowicie okazało się, że prawdopodobieństwo podjęcia zadania spada w wypadku zadań trudnych, a rośnie dla zadań zamkniętych.

Dla analiz najważniejsze jest oszacowanie efektu losowego związanego ze wskaźnikiem motywacji testowej. Okazuje się, że efekt ten jest istotny statystycznie, co oznacza, że wskaźnik znacząco różnicuje osoby badane.

Dla każdej osoby badanej oszacowano wskaźnik motywacji testowej i tak utworzoną zmienną włączono do znanego już nam modelu strukturalnego. Model ten stał się podstawą szacowania interesujących nas zależności.



Schemat 2. Założony układ zależności w modelowaniu strukturalnym (SEM)

Dla prowadzonej analizy kluczowe są wartości współczynników ścieżkowych pomiędzy ogólną motywacją do nauki przedmiotów przyrodniczych a motywacją testową oraz pomiędzy motywacją testową a wynikiem testu PISA 2006 w zakresie przedmiotów przyrodniczych. Odpowiednie analizy SEM przeprowadzono dla wszystkich krajów PISA 2006.

Ze względu na ograniczoną objętość tekstu, ale również z powodu niewielkiej liczby statystycznie istotnych zależności, przedstawione zostaną tylko wybrane wyniki.

Znaczący statystycznie związek ($p < 0,05$) między motywacją ogólną a motywacją testową stwierdzono jedynie w Portugalii. Współczynnik ścieżkowy wyniósł 0,03.

Znaczący statystycznie związek między motywacją testową a wynikiem testu stwierdzono w 5 krajach, jednak nie są to wyniki spójne. W trzech krajach zanotowano dodatnie współczynniki ścieżkowe. W Liechtensteinie współczynnik wyniósł 0,16, w Kolumbii 0,05 i w Norwegii 0,03. W dwóch krajach odnotowano negatywne zależności: w Serbii -0,03, w Grecji -0,04.

Otrzymane wyniki – jak zwykle w wypadku negatywnego rezultatu – można interpretować na dwa sposoby. Pierwsza interpretacja, faktualna, głosiłaby brak znaczących związków między ogólną motywacją do nauki danego przedmiotu a motywacją testową w teście, jeżeli nie ma on znaczących konsekwencji dla rozwiązującego. Równocześnie interpretacja ta wskazywałaby na brak silnego, uniwersalnego wzorca zależności między motywacją testową a wynikiem testu typu *low-stake*.

Druga interpretacja wskazywałaby na wątpliwości związane z trafnością wskaźnika motywacji testowej.

Podsumowanie

Podjęte analizy można osadzić w kontekście problemu teoretycznej trafności testów o niewielkim znaczeniu dla rozwiązujących. Trafność teoretyczna wymaga, by dobre narzędzie miało stabilny, uniwersalny zestaw korelatów dających się wywieść z wiedzy o uwarunkowaniach wyniku testowego. Jeżeli z tej perspektywy spojrzeć na otrzymane wyniki, można pokusić się o dwa główne wnioski. Wyniki testów typu *low-stake* (a przynajmniej testu PISA 2006 w zakresie przedmiotów przyrodniczych) w różnych krajach w odmienny sposób są powiązane z ogólną motywacją do nauki. Dwa wyróżnione wzorce zależności pomiędzy motywacją, statusem rodziny ucznia a wynikiem testu wskazują na znaczące odmienności w psychospołecznych mechanizmach uczenia się w badanych krajach. Ale w świetle pojęcia trafności teoretycznej testu można ten wynik odczytać jako dowód nieadekwatności pomiaru w krajach, gdzie zanotowano bliski zeru (lub nawet ujemny!) związek między motywacją a wynikiem testu.

Drugi wniosek dotyczy motywacji testowej. W tym wypadku zagrożenie dla trafności pomiaru jest bardziej oczywiste, ale w naszych analizach nie znalazło potwierdzenia. Choć trudno w tym wypadku o konkluzywne wnioski. Potrzeba dalszych poszukiwań wskaźników motywacji testowej opartych nie na danych samoopisowych, a obiektywnych pomiarach.

Bibliografia:

1. Atkinson J., W. (1964) An introduction to motivation. Princeton, N. Y.: Van Nostrand.
2. Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, Vol. 37, 122-147.
3. Baumert J., Demmrich A. (2001) Test motivation in the assessment of student skills: The effects of incentives on motivation and performance. *European Journal of Psychology of Education*. Vol. 16, 441-462.
4. Cole J., S., Bergin D., A., Whittaker (2008) Predicting student achievement for low stakes tests with effort and task value. *Contemporary Educational Psychology*. Vol. 33, 609-624.
5. DeMars Ch. (1999) Does the relationship between motivation and performance differ with ability? Paper presented at the Annual Meeting of the National Council on Measurement in Education.
6. Dembo M., H. (1997) *Stosowana psychologia wychowawcza*. Warszawa, WSiP.
7. Eklöf H. (2007) Test-motivation and mathematics performance in TIMSS 2003. *International Journal of Testing*. Vol. 7, No. 3, 311-326.
8. PISA 2006 Technical Report (2007), OECD.
9. Sanchez R., J., Truxillo D., M., Bauer T. N. (2000) Development and examination of expectancy-based measure of test-taking motivation. *Journal of Applied Psychology*, Vol. 85, No. 5, 739-750.
10. Sundre D., L. (1999) Does examinee motivation moderate the relationship between test consequences and test performance? Paper presented at the annual meeting of the American Educational Research Association, Montreal (dostępny w ERIC).
11. Sundre D., L., Wise S., L. (2003) 'Motivation filtering': An exploration of the impact of low examinee motivation on the psychometric quality of tests. Paper presented at the annual meeting of National Council on Measurement in Education, Chicago.
12. Wigfield A., Eccles J. (2000) Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology*.
13. Wise S., L., DeMars Ch. E. (2005) Low examinee effort in low-stake assessment: problems and potential solutions. *Educational Assessment*. Vol. 10, No. 1, 1-17.