

prof. zw. dr hab. Bolesław Niemierko

Akademia Marynarki Wojennej w Gdyni

prof. dr hab. Stanisław Walukiewicz

Instytut Badań Systemowych PAN w Warszawie

Jak mierzyć kapitał ludzki? Ewaluacyjne perspektywy operacjonalizacji pojęć ekonomicznych

*EWD w obecnej postaci to, co najwyżej, egzaminacyjna wartość dodana.
Nasza szkoła prowadzi znacznie szerszą działalność edukacyjną.*

(W. Kossakowski, dyrektor III LO w Gdyni)

Autorzy przedstawiają **kapitał ludzki** i **kapitał społeczny** jako dyspozycje jednostkowe i grupowe będące wynikiem edukacji i mające określoną wartość na rynku pracy. By zbudować narzędzia pomiaru tych form kapitału, obydwa pojęcia należy poddać **operacjonalizacji**, to jest nadać im znaczenie **czynności**, czyli celowych zachowań człowieka. W artykule skoncentrowano się tylko na kapitale ludzkim i omówiono dwie metody operacjonalizacji tego pojęcia: **dwuparametrowy model logistyczny teorii wyniku zadania (IRT)** i **metodę integrującą (MI)**. **Trafność** i **rzetelność** takiego pomiaru kapitału ludzkiego jest rozważana w artykule. Wnioski z analizy prowadzą do rewizji zasad **oceny szkolnego** oraz do poszerzenia pojęcia **ewaluacji w edukacji**.

Pedagodzy włączyli „kapitał” do skarbcza pojęć, którymi opisują i wartościują interesujące ich zjawiska edukacyjne. W Polsce pewien wpływ na to miał zapewne potężny program operacyjny Unii Europejskiej zatytułowany *Kapitał Ludzki. Narodowa strategia spójności*. Zdecydowały jednak o tym nie strategie integracyjno-finansowe, lecz alarmujące przykłady niedostatku kapitału społecznego w szkole (Dudzikowa, 2008) i w środowisku (Marzec-Holka, 2009).

Jako kategoria opisowa, „kapitał” ma duże walory perswazyjne. Za pomocą tego pojęcia bogactwo indywidualnej kompetencji emocjonalnej i poznawczej jest sugestywnie przeciwstawiane ubóstwu odpowiedniego wykształcenia. Jeszcze bardziej wyraziste są porównania zbiorowe. Funkcjonalna sieć związków w pewnej grupie społecznej łatwo przeważa nad rozproszeniem, zubożeniem i marginalizacją członków innej grupy. Wszyscy chcemy więc pomnażać kapitał, zwłaszcza taki, który może posłużyć wyrównywaniu szans między ludźmi.

Edukacja jest dość zgodnie uważana za główny czynnik wzrostu **kapitału niematerialnego**, jak nazywamy łącznie kapitał ludzki i społeczny, a pośrednio – dobrostanu społeczeństwa. To zobowiązuje pedagogów do zajmowania się tą dziedziną. Mogą to czynić na dwa sposoby:

1. **Opisowo**, to jest do odróżniania zjawisk, ich charakteryzowania i sygnalizowania doniosłości, a także do alarmowania opinii publicznej. Podobną rolę odgrywają w edukacji terminy „przemoc symboliczna” i „transmisja kulturowa”, a nawet „mierzenie jakości pracy szkoły”, szczęśliwie już zastąpione w nomenklaturze ministerialnej przez „ewaluację”. Każdy z nich daleko wybiega poza ścisłe procedury, co ułatwia pedagogom kontakt z filozofią i socjologią, ale utrudnia – ze względu na brak ścisłości – współdziałanie z psychologią i z innymi dyscyplinami zorientowanymi eksperymentalnie.
2. **Pomiarowo**, to jest do skalowania wartości indywidualnych i grupowych. Takie podejście otwiera drogę do konwersji kapitału materialnego na rdzenną, **finansową** formę kapitału, do analiz wzajemnej zależności kapitałów ludzkiego i społecznego oraz do badania ich zależności od kontekstu, w jakim są wytwarzane, a w szczególności – od programów edukacyjnych. Konwersja pozwala na niektóre kalkulacje ekonomiczne, ale zrywa więź z humanistyką, dla której człowiek jest nie tylko wartością najwyższą, lecz także z założenia niewymierną.

Widmo zimnej kalkulacji subtelnych spraw ludzkich mogłoby powstrzymać pedagogów od wykraczania poza opisowe, a w istocie – metaforyczne pojmowanie kapitału niematerialnego. Oznaczałoby to jednak rezygnację z dużej szansy przyspieszenia rozwoju tej dyscypliny naukowej.

Szacując kapitał, klasyfikujemy instytucje i osoby na bogate i biedne, co budzi moralny sprzeciw. Podział ludzi na zdrowych i chorych także może być zakwestionowany z pozycji humanistycznych. Gdyby jednak lekarz zrezygnował ze wskaźników stanu zdrowia i nie stosował terapii, w tym interwencji chirurgicznej, życie wielu ludzi przebiegałoby mniej pomyślnie. A gdyby organizacji społeczeństw nie dostosowano do różnic indywidualnych i lokalnych, stracilibyśmy warunki zaspokojenia potrzeb. Rzecz w tym, by klasyfikacji nie rozciągać poza właściwe dla nich cele: leczenia i wspomagania, a unikać **wartościowania uniwersalnego**, bezwzględного wartościowania.

Celem, który chcemy osiągnąć zajmując się kapitałem niematerialnym, jest zaprojektowanie edukacji skutecznej bardziej niż dotychczas. Postęp w mierzeniu dwu form tego kapitału może być bowiem wykorzystany:

1. do badania uwarunkowań wzrostu kapitału niematerialnego,
2. do tworzenia odpowiednich metodyk kształcenia,
3. do ewaluacji jakości edukacji.

Na tak duże korzyści musimy jeszcze poczekać. Pierwszym etapem działania będzie nadanie pojęciom „kapitału ludzkiego” i „kapitału społecznego” formy pozwalającej określić ich zasób w obiektach rzeczywistych, nie zaś tylko idealnych, w których jest on z założenia pełny. By utrzymać ten referat w rozsądnych granicach, ograniczymy się do kapitału ludzkiego.

Dobór metod szacowania właściwości ludzi jest zadaniem psychometrii i „edumetrii”, jak się czasem nazywa pomiar pedagogiczny. Dzięki tym metodom możemy uzyskać dyrektywy wspomagające rozwój kapitału niematerialnego zarówno w skali wąskiej, jednostkowej, jak i w skali szerokiej, systemowej.

1. Pojęcia kapitału ludzkiego i kapitału społecznego

Pojęcia z zakresu kapitału niematerialnego są tak różnorodne i tak w naukach społecznych rozmyte, że ich przegląd by zagmatwał podjęte zagadnienie i nie pozwolił na zaprojektowanie pomiaru tego kapitału. Inaczej więc niż w typowych rozprawach na ten temat (Domański, 1993; Kwiatkowski, 2005), za punkt wyjścia objerzemy nie mnogie i rozbudowane, lecz pojedyncze i zwięzłe definicje. Znajdziemy je w podręcznikach Stanisława Walukiewicza (2010; *w przyg.*), wydawanych w Instytucie Badań Systemowych PAN w ramach projektu współfinansowanego ze środków UE w ramach Europejskiego Funduszu Społecznego *Kapitał ludzki i kapitał społeczny jako nowe przedmioty akademickie*.

Zacznijmy od pojęcia **kapitału** jako *zasobu, który daje lub może dać zysk* (Walukiewicz, 2010, s. 24). W tej lapidarnej definicji akcent pada na „zysk”, jako nadwyżkę przychodów nad kosztami. Ta nadwyżka nie jest ograniczona do korzyści finansowej, lecz może także przybrać formę materialną (przedmioty fizyczne) lub niematerialną (zaspokojenie potrzeb psychicznych człowieka).

Jak głosi definicja, kapitał *daje* lub *może dać* zysk. Ta alternatywa jest istotna, bo dla edukacji zasadnicza jest nie aktualna, lecz potencjalna realizacja zysku. Pedagog pracuje dla przyszłości (Suchodolski, 1947). Ani w skali jednostki, ani w skali masowej nie ma gwarancji, że okoliczności pozwolą na zysk z wybranego elementu wykształcenia. Niepewność jest tu większa niż w gospodarce, co powinno nas przestrzec przed jej dalszym zwiększaniem przez błędy pomiaru.

Przejdźmy teraz do kluczowego dla nas pojęcia „kapitału ludzkiego”. W cytowanym podręczniku (tamże, s. 25) czytamy: **Kapitał ludzki** to *wszystkie, ale to absolutnie wszystkie, zasoby niematerialne (przymioty) kojarzone z człowiekiem traktowanym jako samodzielna istota ludzka*.

Także i ta definicja wymaga kilku komentarzy:

1. „Wszystkie, ale to absolutnie wszystkie” – jak podkreśla jej autor – zasoby niematerialne człowieka wymagają (1) ewidencji tych zasobów, (2) pomiaru każdego elementu z osobna lub, co bardziej realne, w dobrze wyważonych skupieniach (aglomeratach). Jest zrozumiałe, że zaczynamy od zasobów najważniejszych, wokół których można zgrupować pozostałe.
2. Przymioty „kojarzone z człowiekiem” są odwołaniem do osądu zewnętrznego, do eksperta, który dokonuje skojarzeń, posługując się procedurami. Podmiot nie musi zdawać sobie sprawy z wielkości swoich walorów ani nawet z ich istnienia. Na przykład, uczeń może nie wiedzieć, że ma talent, a jedną z ról nauczyciela jest ten talent odkryć i go rozwijać.
3. Kapitał ludzki jest indywidualny. Gdy jest dzielony z innym człowiekiem, w sensie jednostronnego lub wzajemnego uzależnienia, należy do kategorii społecznej. Kapitałem ludzkim jest to, co jednostka „nosi w sobie” samodzielnie i może zademonstrować w różnych sytuacjach materialnych i układach personalnych oraz w odpowiednich testach.

Autor tej definicji nie kusi się o pełne skatalogowanie zasobów kapitału ludzkiego. Jako przykłady podaje pięć grup właściwości człowieka:

1. kompetencje i doświadczenie,
2. wiedza i zdolności (intelektualne, artystyczne, przywódcze itp.),
3. zdrowie i wydolność organizmu,
4. nastawienie życiowe (optymizm – pesymizm, otwartość – skrytość, prawdomówność – hipokryzja itp.),
5. zadatki na kapitał społeczny (zaufanie – podejrzliwość, tolerancja – kłótność, lojalność – nielojalność itp.).

Dyskusję tych składników kapitału ludzkiego podejmiemy w następnej części referatu, po zaprezentowaniu pojęcia „kapitału społecznego”. Brzmi ona (Walukiewicz, w *przygotowaniu*, s. 8): **Kapitał społeczny** to *formalne i/lub nieformalne relacje między co najmniej dwiema osobami traktowanymi jako samodzielne istoty ludzkie*.

Ta definicja zawiera elementy wspólne z poprzednią, a są to „samodzielne istoty ludzkie” i odpowiednie „traktowanie” ich przez szacującego kapitał. Nowy element to **relacje**, rozumiane jako stosunki między ludźmi oparte, według Piotra Sztompki (2002, s. 8), na (1) **więzi obiektywnej**, wynikającej z podobieństwa sytuacji życiowej, na (2) **więzi kooperacyjnej**, wynikającej z realizacji wspólnych celów, oraz na (3) **więzi moralnej** – zaufaniu, lojalności i solidarności. Podobny repertuar więzi wymienia Walukiewicz (2007), rozróżniając cztery *formy bliskości*: przestrzenną, organizacyjną, emocjonalną i twórczą.

2. Operacjonalizacja pojęć pedagogicznych

Pedagogika, jako nauka dyrektywalna, teoretyczno-praktyczna (Muszyński, 1971), jest szczególnie narażona na **myślenie życzeniowe**, zmieniające obraz rzeczywistości na zgodny z oczekiwaniami. Łatwo przyswajają sobie terminologię innych dyscyplin, zwłaszcza tę, która dotyczy właściwości jednostek i grup społecznych. To dotyczy emocji, postaw, umiejętności i dziedzin wiedzy, których opanowanie jest pożądane, ale niełatwe do zaobserwowania. Chętnie łudzimy się, że nasycenie programów kształcenia odpowiednią terminologią zapewni nam sukces edukacyjny.

Około pół wieku temu podjęto w Stanach Zjednoczonych walkę z mętym formułowaniem celów kształcenia, nalegając, by zawsze odnosiły się do zachowań (*behavioral objectives*). Klasykiem tego kierunku stał się Robert Mager (1962; 1972; 1990), obdarzony błyskotliwym humorem, przydatnym w tępieniu ogólnikowych, a zarazem patetycznych sformułowań, nienadających się do pomiaru osiągnięć uczniów.

Przekształcenie pojęcia z postaci **ogólnej**, abstrakcyjnej i hasłowej, a bywa że sloganowej, na postać **operacyjną**, umożliwiającą rozpoznawanie określonej klasy faktów, nazywamy jego **operacjonalizacją**. Dorobek autorski Magera można ująć w następującą procedurę (Niemierko, 1999, s. 41):

1. Zapisanie celu w postaci ogólnej, np. *rozwój osobowości ucznia*.
2. Intuicyjny obraz ucznia osiągającego cel. Jak uczeń o „rozwinętej osobowości” zachowuje się w określonych sytuacjach? Czym to zachowanie różni się od zachowania ucznia, który celu nie osiągnął?

3. Luźne zapisy (*jottings*) celów operacyjnych, uzyskiwane w toku indywidualnej lub zbiorowej „burzy mózgów”, np. *docenia swoje walory, przejawia optymizm, ma realistyczne plany życiowe, podejmuje przemyślane decyzje*.
4. Selekcja luźnych zapisów. Zadaniem selekcji jest wyeliminowanie lub przeredagowanie tych zapisów, które są jeszcze nie dość operacyjne, gdyż nie opisują obserwowalnej czynności ucznia, np. „docenia swoje walory” warto zmienić na *trafnie rozróżnia swoje mocne i słabe strony oraz w sytuacjach trudnych zachowuje dodatnią samoocenę*.
5. Klasyfikacja luźnych zapisów. Zapisy czynności należy pogrupować według stosownego modelu teoretycznego, np. hierarchicznie, wykorzystując odpowiednią taksonomię celów emocjonalnych kształcenia: od chętnego uczestnictwa w działaniach społecznych do swoistego, dojrzałego systemu działań (Niemierko, 2007, s. 93-96). Ten zabieg ułatwi konstrukcję narzędzi pomiaru.
6. Sformułowanie celów operacyjnych. Budujemy zdania, w których podmiotem jest uczeń („Uczeń potrafi...”), starając się w nich zawrzeć jak najwięcej uznanych i poklasyfikowanych luźnych zapisów. Jednym z celów operacyjnych rozwoju osobowości ucznia mogłoby stanowić sformułowanie: *W sytuacjach trudnych uczeń wykazuje opanowanie, wytwarza własny plan działania i realizuje go, do końca zachowując przekonanie o możliwości odniesienia sukcesu*.
7. Sprawdzenie celów operacyjnych. Trafność i operacyjność naszej listy celów sprawdzamy przez zaplanowanie sytuacji sprawdzania, czy cele operacyjne są osiągnięte, a pośrednio – czy osiągnięty jest cel ogólny, a następnie zbudowanie arkusza obserwacji, kwestionariusza lub testu, zastosowania tego narzędzia i staranną analizę wyników.
8. Ewentualne powtórzenie etapów 2 – 7. Gdy wynik sprawdzenia jest niezadowolający, powtarzamy etapy operacjonalizacji celu, usuwając jej błędy i ograniczenia. Procedura operacjonalizacji celów jest bowiem *iteracyjna*, prowadzi do pożądanego wyniku przez kolejne przybliżenia.

Operacjonalizacja sprowadza pojęcia teoretyczne „na ziemię” przez dobór **wskaźników empirycznych** występowania zjawisk, w tym osiągnięcia celów kształcenia przez uczniów. Jest niezbędna do wiązania ideałów humanistyki z oporną rzeczywistością (Pawłowski, 1969). Nie jest wolna od subiektywizmu, bo opiera się na indywidualnym lub grupowym wyobrażeniu desygnatu pojęcia (kroki 2. i 3. procedury), ale umożliwia *pomiar* zjawisk, jako rejestrowanie przypadków ich występowania i ich **ewaluację**, jako szacowanie wartości.

Warto przypomnieć, że pionierzy operacjonalizacji celów kształcenia próbowali ją rozciągnąć na klasy zadań testowych („druga operacjonalizacja”), a nawet na gotowe narzędzia pomiaru („trzecia operacjonalizacja”), co daleko wykraczało poza precyzowanie pojęć (Schott, 1972, s. 69). Wkrótce jednak utożsamienie operacjonalizacji celów z konstrukcją testu, nazwane przez wybitnego teoretyka pomiaru Lee Cronbacha „ultraoperacjonalizmem” (1969), zostało zakwestionowane i kontrola posługiwania się celami pomiaru pozostała zagadnieniem **trafności teoretycznej** pomiaru, czyli zgodności jego procedur i wyników z wybranymi teoriami psychologii i pedagogiki.

3. Trafność pomiaru kapitału ludzkiego

Trafność pomiaru, pojmowana jako *stopień, w jakim dane empiryczne oraz teoria wspierają interpretacje wyników w proponowanych zastosowaniach narzędzi pomiaru* (Standards..., 1999, s. 9), a w uproszczeniu – potwierdzona użyteczność jego wyników do pewnego celu, jest w naukach społecznych zgodnie uważana za kardynalną właściwość pomiaru. Kłopoty z trafnością powodują, że wielu humanistów uważa przymioty ludzkie za z gruntu niemierzalne, a uściślonę procedury ich analiz za naiwny „pozytywizm”. Szanujemy te poglądy, ale do celu ewaluacji w edukacji musimy znaleźć mocniejsze empirycznie rozwiązanie.

Spośród wielu odmian trafności i stanowisk metodologicznych w tym zakresie wybieramy koncepcję Samuela Messicka (1980) jako najogólniejszą, powszechnie uznawaną, a nadto – co ważne dla pedagoga – nachyloną etycznie. Uważał on, że podstawowe pytania, na które powinien sobie odpowiedzieć twórca procedury pomiarowej w zakresie nauk społecznych, dadzą się usystematyzować jak w tab. 1.

Tabela 1. Trafność teoretyczna pomiaru według Samuela Messicka

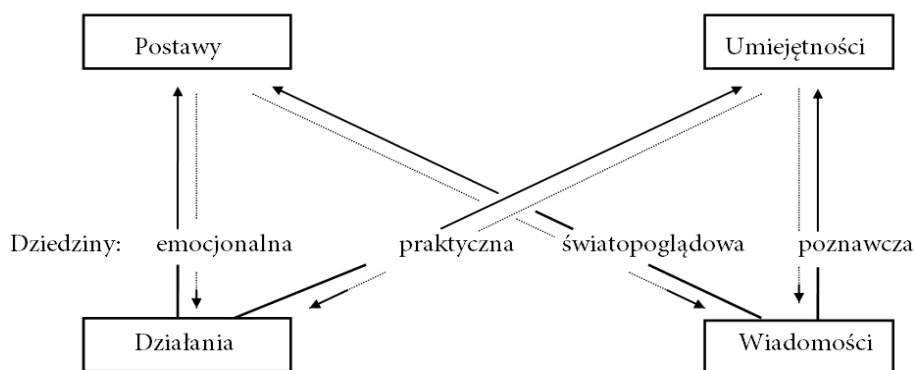
Podstawa wnioskowania	Interpretacja wyniku	Wykorzystanie wyniku
Uzyskane dowody	1. Interpretacja pojęć <i>Czy zebrany materiał upoważnia do określonej interpretacji?</i>	2. Praktyczna użyteczność <i>Czy decyzje podjęte na podstawie pomiaru są uzasadnione?</i>
Przewidywane konsekwencje	3. Wynikanie wartości <i>Czy wynikające wartości są tymi, które należy upowszechniać?</i>	4. Skutki społeczne <i>Czy zamierzone i niezamierzone skutki pomiaru będą pomyślnie?</i>

Tak szeroko pojmowana trafność stawia przed stosującymi go wysokie wymagania, zwłaszcza co do przewidywanych konsekwencji, nazwane potem **trafnością konsekwencyjną** pomiaru. Wprowadzający nowy termin i nową procedurę na rynek edukacyjny muszą dołożyć starań, by zapobiec ich nadużyciom oraz skutkom ubocznym, które mogą zniweczyć wszelką korzyść z innowacji. W przypadku kapitału ludzkiego zagrożenia są następujące (numeracja według tab. 1.):

1. Zbyt wąskie traktowanie tego kapitału: tylko jako wiadomości i umiejętności, tylko jako kwalifikacje zawodowe. Takie zubożenie zakresu Messick (1989, s. 34-35) nazwał **niedoreprezentowaniem pojęcia** (*construct underrepresentation*). Może być ono powodowane ograniczeniem zadania poznawczego (przedmiot szkolny, rekrutacja pracowników) lub wygodnictwem diagnostów (wybór cech najłatwiej sprawdzalnych). Równie często pojawia się **wariancja niestosowna** (*construct-irrelevant variance*) jako skażenie (*contamination*) wyników z przyczyn metodologicznych: wahań dyspozycji i motywacji badanych, różnic kulturowych między badanymi, wieloznaczności zadań i pytań, subiektywizmu obserwatorów oraz podobnych czynników wymykających się kontroli.

2. Wartość kapitału ludzkiego silnie zależy od sytuacji, w której będzie wykorzystywany. **Sytuacja** jest układem bodźców wewnętrznych i zewnętrznych działających na człowieka w określonym momencie czasu (Tomaszewski, 1975, s. 17 i n.). Gdy ten układ się zmienia, jego zachowania także ulegają zmianie. Decyzje o uczniu lub pracowniku podejmowane na podstawie diagnozy jego kapitału ludzkiego w sytuacjach *A* i *B* (na przykład w toku zajęć szkolnych i podczas egzaminu) mogą być ekstrapolowane na sytuacje *C* i *D* (w pracy zawodowej i w życiu społecznym) powiązane z poprzednimi teorią naukową i potocznym doświadczeniem. To jest wszakże przejście od *diagnozy* do *prognozy*. Ryzyko prognozy może być duże i trudne do oszacowania.
3. Przesłaniem, jakie oszacowania kapitału ludzkiego niosą dla niespecjalisty, jest **kapitałowa wartość człowieka**: niska dla nieletnich i słabo wykształconych, znacznie wyższa dla dorosłych, zwłaszcza uzdolnionych (biznesowo, intelektualnie, sportowo, artystycznie). Wartością najbardziej godną upowszechnienia w społeczeństwie jest natomiast **praca nad sobą**, wciąż mało popularna wśród naszej młodzieży (Dudzikowa, 1993). Wyniki pomiaru kapitału ludzkiego powinny być zatem interpretowane **perspektywicznie**, co najmniej na dystans pełnej dojrzałości danego pokolenia, i **warunkowo**, dla różnych poziomów własnego zaangażowania osoby.
4. Skutki społeczne pomiaru kapitału ludzkiego trzeba starannie przewidywać i śledzić (monitorować), gdyż w naukach społecznych droga od haseł („rozwijajmy!”) do faktów (wyniki działania) obfituje w niespodzianki. Pojawia się tu **samospełniające się proroctwo** polegające na nieświadomym wytwarzaniu danych potwierdzających opinię (Tyszka, 1999). Na przykład informowanie uczniów o ich ilorazie inteligencji, uważanym za względnie stałą właściwość człowieka, zniechęca młodzież do wysiłku. Ostrożność w posługiwaniu się oszacowaniem kapitału ludzkiego musi być szczególnie wysoka (Niemierko, 2009, s. 41-45). „Przede wszystkim – nie szkodzić” powinno być dewizą nie tylko lekarzy, lecz także pedagogów stosujących diagnozę edukacyjną.

By przystąpić do pomiaru kapitału ludzkiego wytwarzanego w procesie edukacyjnym, musimy stworzyć mapę właściwości składających się na kapitał ucznia i absolwenta szkoły. Wobec braku powszechnie uznanej, ponadbranżowej klasyfikacji ekonomicznej kapitału pracownika sięgniemy do dorobku nauk pedagogicznych w zakresie porządkowania celów kształcenia. Przydatna wersja może być oparta na dorobku Beniamina Blooma (1956) i Dawida Krathwohla (1964) oraz na sformułowanych przez Bolesława Niemierkę czterech **taksonomiach celów kształcenia**: emocjonalnej, poznawczej, praktycznej i światopoglądowej (Niemierko, 1990, s. 230; 2007, s. 110). Tworzą one układ przedstawiony na rys. 1.



Rysunek 1. Cztery dziedziny celów kształcenia

Rys. 1. pokazuje, że zarówno działania (czynności) wykonywane przez uczniów, jak i wiadomości (informacje), jakie do nich docierają, kształtują ich postawy i umiejętności (strzałki „dół – góra”), czemu towarzyszy także, na początek słabsze, oddziaływanie zwrotne (strzałki „góra – dół”). Cztery dziedziny tworzą układ o silnej i wyrazistej strukturze. Jego zewnętrzne filary odpowiadają dwu rodzajom procesów psychicznych obejmowanych kształceniem (Doliński, 2000; Niemierko, 2007, s. 23), a wewnętrzna „kratownica” wiąże te procesy w system.

Adaptacja modelu taksonomii celów kształcenia do problematyki pomiaru kapitału ludzkiego może przybrać postać następującą:

1. **Dziedzina fizyczno-praktyczna, FP.** Tu mieszczą się podstawowe dyspozycje fizyczne: zdrowie i wydolność organizmu, a także sprawności ruchowe (kinestetyjne): samoobsługowe (np. przemieszczanie się), szkolne (np. pisanie ręczne i maszynowe), sportowe, techniczno-artystyczne. Osiągnięcia z tego zakresu wymienia Walukiewicz w trzeciej i w drugiej grupie podanych przykładów.
2. **Dziedzina poznawczo-doświadczalna, PD.** Obejmuje najbardziej znane inteligencje wielorakie: językową, logiczno-matematyczną i przestrzenną (Gardner, 2002), wykształcenie ogólne i ukierunkowane, a także doświadczenie zdobyte poza szkołą, w szczególności – w zakładzie pracy. Osiągnięcia z tego zakresu wymienia Walukiewicz w grupach drugiej i pierwszej podanych przykładów.
3. **Dziedzina emocjonalno-motywacyjna, EM.** Obejmuje reakcje na bodźce otoczenia (o podłożu temperamentalnym), ciekawość poznawczą i chęć uczenia się, zasoby energii i wytrwałość (persewerację). Niektóre z nich znajdziemy wśród przykładów Walukiewicza w grupie czwartej.
4. **Dziedzina charakterologiczno-światopoglądowa, CS.** W tej niedocenianej na ogół kategorii rozwoju indywidualnego mieszczą się: dodatnia samoocena (asertywność) i uspołecznienie, a także należycie rozumiana ideowość, wyrażająca się w postawach wobec świata zewnętrznego: przyrody, społeczeństwa i kultury. Te wartości sygnalizuje Walukiewicz pod hasłami „nastawienia życiowego” i „zadatków na kapitał społeczny” w grupach czwartej i piątej podanych przykładów.

Zarysowana powyżej topografia (mapa) obszaru pojęcia „kapitał ludzki” stanowi co najwyżej punkt wyjścia do badań nad interpretacją tego pojęcia w zmiennej (branżowo, czasowo, organizacyjnie) sytuacji rynku pracy. Te badania nadadzą wagi poszczególnym dziedzinom. Na razie przyjmujemy, że wagi są równe, co jest śmiałym wyzwaniem wobec panujących poglądów, w których zwykle wyolbrzymia się dziedzinę poznawczo-doświadczalną i nie docenia innych dziedzin.

Po uzgodnieniu interpretacji pojęcia kapitał ludzki i po dyskusji wartości, jakie wnosi do edukacji (patrz: lewa kolumna tab. 1.), można będzie ostrożnie spróbować jego zastosowań, bacznie śledząc użyteczność i konsekwencje społeczne każdego kroku (patrz: prawa kolumna tab. 1.). Dopiero wtedy można będzie ocenić trafność projektu operacjonalizacji pojęcia kapitał ludzki.

Podjęto już także pierwsze krajowe prace nad pomiarem kapitału społecznego w edukacji (Wiktorzak, 2009; Wendicki, w *przyg.*), ale pełna operacjonalizacja tego pojęcia, wymagająca syntezy badań z dziedziny pedagogiki, psychologii społecznej, psychologii pracy i socjologii, nie została jeszcze dokonana. Przyczynki pedagogiczne do takiej syntezy przedstawił Andrzej Janowski jako „zjawiska zachodzące w zespole wychowanków” (2002, r. 3).

4. Dwuparametrowy model logistyczny teorii wyniku zadania (IRT)

Pomiar pedagogiczny od wielu lat toczy walkę o wyodrębnienie w **psychometrii**, opartej na skalowaniu różnic indywidualnych między ludźmi. Rozpoczęto ją w epoce ekspansji nauczania programowanego (Glaser, 1961) pod hasłem budowy testów **sprawdzających** opanowanie wybranego zakresu treści kształcenia (*criterion-referenced tests, mastery tests*), a nie tylko **różnicujących** osiągnięcia między uczniami (*norm-referenced tests*). Powstało wiele opracowań z tego zakresu (Niemierko, 1990), a zasady konstrukcji testów sprawdzających zostały upowszechnione w Polsce na setkach kursów i studiów podyplomowych.

Słabością pomiaru sprawdzającego jest chwiejna procedura wyznaczania **norm ilościowych** (*performance standards, cutting scores*), pozwalających na uznanie osiągnięć za wystarczające. Żadna wersja tej procedury nie uwalnia nas od dostosowania normy do stanu osiągnięć uczniów i od rozważenia konsekwencji podejmowanych decyzji (Popham, 1978; Haertel, 2002). Z tego powodu będziemy musieli poczekać z jej proponowaniem do czasu zbudowania narzędzi pomiaru kapitału ludzkiego i kapitału społecznego oraz ich dostatecznie starannego wypróbowania (standaryzacji). Na razie projektujemy skale pomiarowe w konwencji pomiaru różnicującego.

Skale powinny obejmować cały zakres zmienności kapitału ludzkiego: od pojawienia się dyspozycji (narodzin człowieka) do mistrzostwa (poziomu eksperta). Tak długie skale były budowane w pedagogice już w latach osiemdziesiątych ubiegłego stulecia (Larkin i in., 1980; Glaser, 1988) pod hasłem pomiaru określonych kompetencji na kontinuum „nowicjusz – ekspert”. Ich znaczenie wzrosło w kolejnych dziesięcioleciach, gdy teoria wyniku zadania (*Item Responce Theory, IRT*) wzięła górę nad klasyczną teorią testu (Hambleton, 1993). W Polsce dopiero zaczynamy wdrożenia teorii wyniku zadania (Szaleniec, 2010), ale jej przydatność badawcza dla pedagogiki wydaje się przesądzona.

Teoria wyniku zadania (probabilistyczna teoria testu) jest teorią matematyczną dotyczącą związku wyniku zadania z właściwością mierzoną przez zadania danego testu (Hambleton, 1993). Ta właściwość stanowi **zmienną ukrytą**, zdefiniowaną w toku planowania narzędzia, ale nieobserwowalną, przejawiającą się jedynie w postaci wyników poszczególnych zadań (elementów narzędzia), a dokładniej: w postaci prawidłowości układu tych wyników. By badać te prawidłowości, musimy przyjąć mocne założenie o **lokalnej niezależności** wyników poszczególnych zadań, czyli o jednorodności mierzonej właściwości. „Pytanie o to, czy narzędzie pomiarowe jest wystarczająco jednowymiarowe, stanowi podstawowe pytanie umożliwiające stosowanie modeli IRT [teorii wyniku zadania]” – czytamy w podręczniku tej teorii (Hulin i in., 2005, s. 239). Dokładność dopasowania modelu, a zatem i zasadność opartych na nim oszacowań, jest, podobnie jak lokalna niezależność zadań, sprawdzana empirycznie (Hambleton, 1989, s. 172-182).

Założenie jednorodności nie wyklucza pomiaru właściwości wielowymiarowych, jak wyżej scharakteryzowane dziedziny kapitału ludzkiego, ale każdy z wymiarów musi być jednorodny, a jego wewnętrzna spójność musi znaleźć potwierdzenie w układzie wyników zadań. Zadania (pozycje skali), które naruszają ten układ, powinny być wyeliminowane z analiz, co może obejmować nawet większość zadań poddanych próbie jednorodności. Z reguły są to zadania o niskiej *mocy różnicującej*, czyli zdolności rozróżniania badanych o wyższej i niższej wartości całego testu.

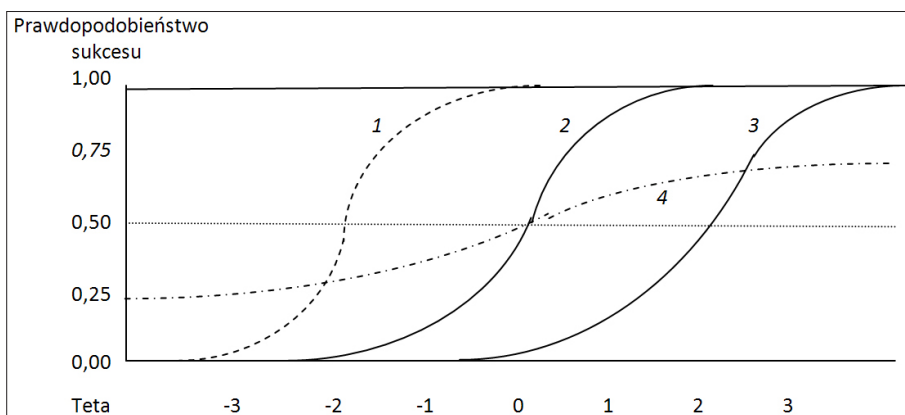
Drugie założenie teorii dotyczy zależności wyniku od trudności zadania i od zdolności badanego, mierzonych pozycją na skali zmiennej ukrytej, a więc – na wspólnej skali. Gdy te pozycje są jednakowe, badany ma prawdopodobieństwo rozwiązania zadania równe 0,50, gdy pozycja zadania jest wyższa, to prawdopodobieństwo jest mniejsze, a gdy pozycja badanego jest wyższa – większe.

Najbardziej dla nas przydatna **dwuparametrowa funkcja logistyczna** zmiennej ukrytej (*model 2PL*) przybiera postać następującą (Verhelst, 2010, s. 44):

$$f_i(\theta) = \frac{\exp[a_i(\theta - b_i)]}{1 + \exp[a_i(\theta - b_i)]},$$

gdzie θ oznacza zmienną ukrytą „teta”; \exp – wykładnik stałej matematycznej $e = 2,7182\dots$, będącej podstawą logarytmu naturalnego; a_i – parametr różnicowania zadania i ; b_i – parametr trudności zadania i .

Graficzne przedstawienie funkcji $f_i(\theta)$ nazywamy **funkcją charakterystyczną zadania i** . Cztery krzywe charakterystyczne zadań 1, 2, 3 i 4, o różnych parametrach a i b , widzimy na rys. 2.



Rys. 2. Cztery krzywe charakterystyczne zadań testowych

Parametr a , **różnicowanie**, jest nachyleniem (stromością) krzywej charakterystycznej w punkcie przegięcia między częścią wklęsłą a wypukłą, wyznaczonym przez parametr b . Może przybierać wartości od nieskończenie niskich do nieskończenie wysokich, ale w praktyce mieści się najczęściej między 0,5 a 2,5 (Hornowska, 2001, s. 198). Na rys. 2. różnicowanie przybrało w przybliżeniu następujące wielkości: $a_1 = 2,5$ (zadanie bardzo dobrze różnicujące); $a_2 = 2,0$ (zadanie dobrze różnicujące); $a_3 = 1,5$ (zadanie średnio różnicujące); $a_4 = 0,5$ (zadanie słabo różnicujące). Znaczenie różnicowania jest dla teorii wyniku zadania zasadnicze: im jest wyższe, tym więcej mamy informacji o badanych uzyskujących wyniki w zasięgu różnicowania przez dane zadanie.

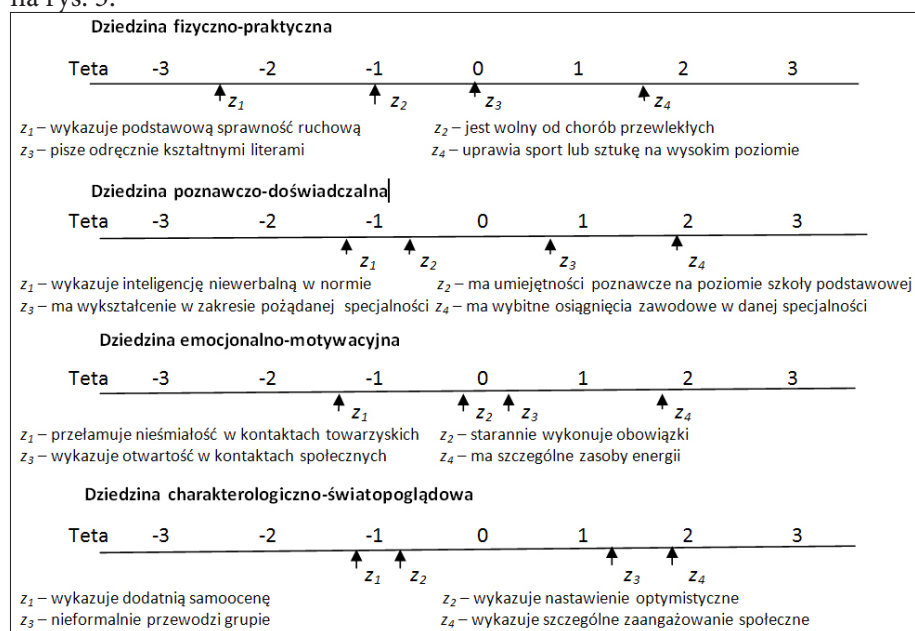
Parametr b , **trudność**, jest położeniem na **skali zdolności** badanych (teta) punktu, w którym prawdopodobieństwo rozwiązania zadania wynosi 0,50. Także i ten parametr może przybierać wartości od nieskończenie niskich do nieskończenie wysokich, ale przedział od -3,00 do 3,00 zwykle wystarcza, by przedstawić jego wartości. Przeciwnie niż w przypadku wskaźnika łatwości (p) w klasycznej teorii testu, gdy wartość parametru b rośnie, to łatwość zadania maleje. Trudność zadań przedstawionych na rys. 2. jest w przybliżeniu następująca: $b_1 = -2,0$ (zadanie bardzo łatwe); $b_2 = 0,0$ (zadanie umiarkowanie trudne); $b_3 = 2,0$ (zadanie bardzo trudne); $b_4 = 0,0$ (zadanie umiarkowanie trudne).

Teoria wyniku zadania może być zastosowana nie tylko do zadań testowych, lecz także do pozycji kwestionariuszowych (skale ocen i samooceny, skale postaw, skale opisowe), kategorii zachowań z arkuszy obserwacji, a nawet do informacji uzyskiwanej w analizie dokumentów. Warunkiem należytej rzetelności, czyli powtarzalności wyniku pomiaru, jest zbudowanie dostatecznie jednorodnej (homogenicznej) skali z dostatecznej liczby wysoko różnicujących elementów.

Inaczej niż w klasycznej teorii testu, w których oszacowanie rzetelności jest jedno dla całej długości skali, teoria wyniku zadania umożliwia dostosowanie rzetelności testu do grupy badanych lub do pojedynczego badanego. Na podstawie parametrów zadania można wykreślić nie tylko jego krzywą charakterystyczną, lecz także **krzywą informacyjną zadania**, pokazującą, ile informacji

dostarcza to zadanie w każdym punkcie skali zdolności teta. Ta krzywa przybiera postać rozkładu normalnego o wierzchołku w punkcie parametru b , wysokości (rzędnej) proporcjonalnej do parametru a i zasięgu ramion obejmującym obszar wzrostu krzywej. Informacja dostarczana przez poszczególne zadania sumuje się dla testu złożonego z tych zadań, co pozwala na wymodelowanie **krzywej informacyjnej testu** według potrzeb diagnostycznych grupy. Jeszcze większe znaczenie ma to w **testowaniu adaptacyjnym**, gdzie pozwala na dobieranie zadań, które dostarczą najwięcej informacji o badanym według bieżącego oszacowania jego poziomu zdolności (teta) lub według poziomu oczekiwań zakładu pracy wobec kandydatów na określone stanowiska.

Dla ilustracji skal teta czterech dziedzin kapitału ludzkiego zaznaczymy po cztery hipotetyczne punkty (zadania: z_1, z_2, z_3, z_4) na każdym kontinuum, jak na rys. 3.



Rysunek 3. Hipotetyczny przykład skalowania składników kapitału ludzkiego

Rys. 3. daje nam orientację co do wskaźników (zadań), jakie mogą być zastosowane w procesie operacjonalizacji i skalowania czterech dziedzin kapitału ludzkiego. By przejść od nich do pomiaru osiągnięć w tych dziedzinach, trzeba:

1. dobrać po kilkadziesiąt (co najmniej po 30 – 40) wskaźników (elementów narzędzia) do każdej skali,
2. zebrać dane empiryczne w reprezentatywnej próbie (o liczności rzędu 1000 – 2000) osób z odpowiednio zdefiniowanej populacji,
3. wyeliminować wskaźniki o niskiej mocy różnicującej (niskiej wartości parametru a),
4. oszacować wartość skalową teta (parametr b) pozostałych wskaźników,

5. oszacować rzetelność pomiaru i standardowy błąd pomiaru w kolejnych przedziałach skali teta,
6. określić trafność teoretyczną, w tym trafność konsekwencyjną pomiaru o danym poziomie rzetelności,
7. przygotować podręcznik pomiaru kapitału ludzkiego do celu odpowiednio ostrożnych wdrożeń opisanej metody.

Te siedem kroków mogłoby być wieloletnim programem całego instytutu badawczego lub przynajmniej dużego seminarium naukowego. Gdyby nawet takiemu zespołowi nie udało się uzyskać skal pomiarowych o zadowalającej trafności i rzetelności, to doświadczenie zdobyte w toku badań powinno zawoalować postępowaniem w pedagogice i w innych naukach społecznych.

5. Metoda integrująca

Zastosowanie modelu logistycznego w psychologii i pedagogice polega na **dedukcji**, na przechodzeniu od abstrakcyjnych założeń dotyczących zmiennych ukrytych do zachowań badanych osób. Wymaga żmudnego poszukiwania empirycznych wskaźników, które pozostawałyby w zgodzie z apriorycznym modelem układu tych zmiennych.

Podejściem mniej ryzykownym i mniej pracochłonnym jest **metoda integrująca**, polegająca na wykorzystaniu ogólnie dostępnej informacji o zachowaniach badanych osób do wnioskowania o ogólnych prawidłowościach tych zachowań, a więc na **indukcji**. Im szerszy jest zakres integracji danych, **poziomy** – co do ilości informacji zbieranej o każdym z badanych i **pionowy** – co do liczby etapów działania edukacyjnego, tym większa jest wartość poznawcza diagnozy.

Największym zbiorem danych o rozwoju kapitału ludzkiego przez edukację są w Polsce wyniki kolejnych egzaminów zewnętrznych. Najprostszym zastosowaniem metody integrującej jest więc wykorzystanie wyników egzaminów, jakie przechodzi dany uczeń/student, powiedzmy Janek Kowalski (JK), w procesie swej edukacji „od domu/przedszkola/szkoły podstawowej do dyplomu licencjata/magistra/doktora” (Walukiewicz, 2010, s. 137). Czekają go na tej drodze trzy egzaminy doniosłe (po szkole podstawowej, gimnazjum i liceum), ale można też zebrać dane z badań gotowości szkolnej (Jasińska, 2010), w przygotowaniu są sprawdziany po trzech klasach szkoły podstawowej, a egzaminy dyplomowe w szkołach wyższych też zawierają pewne elementy standaryzacji.

Nie ulega wątpliwości, że edukacja jest jednym z wielu sektorów gospodarki rynkowej, więc coraz częściej jest stawiane pytanie o efektywność inwestycji w tym sektorze. Na przykład rodzice JK mogą zadać pytanie, na ile pieniądze z ich podatków, które poszły na edukację, przyczyniły się do wzrostu wartości kapitału ludzkiego ich syna. By odpowiedzieć na to pytanie, musimy znaleźć metodę pomiaru zmian wartości kapitału ludzkiego ucznia/studenta w procesie edukacyjnym.

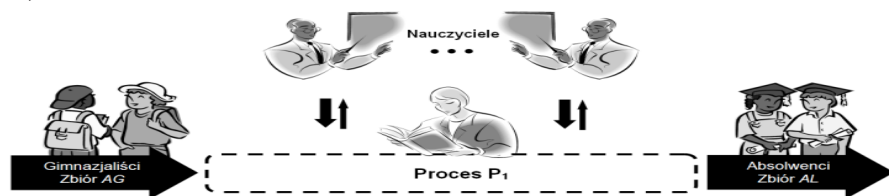
Prób pomiaru wzrostu osiągnięć uczniów dokonano w Polsce już wiele, głównie w gimnazjach, w których wypróbowano metodę **edukacyjnej wartości dodanej** (EWD), polegającą na szacowaniu przyrostu osiągnięć uczniów: **względny** – w jednostkach standardowych (Dolata, 2008) lub **bezwzględny** – w skali

równoważników klasy (Niemierko, 2009, s. 254-259 i 283-286). Obydwa podejścia zaowocowały wykryciem doniosłych zależności (postępująca segregacja kandydatów do gimnazjów wielkomiejskich, koszt psychologiczny sukcesu dydaktycznego), a Centralna Komisja Egzaminacyjna i Instytut Badań Edukacyjnych kontynuują pracę w tym zakresie (Lisiecka, 2010). Pojawiają się też pierwsze projekty rozciągnięcia oszacowań edukacyjnej wartości dodanej na cały okres kształcenia ogólnego (w amerykańskiej nomenklaturze: *K – 12*).

Obecne egzaminy końcowe w szkole podstawowej (sprawdzian) i średniej (egzamin gimnazjalny i matura) są ujednoczonymi egzaminami zewnętrznymi, w których uczniowie mają do rozwiązania te same zadania. Nastręcza to pewne trudności organizacyjne, ale ułatwia porównania poziome osiągnięć (ich różnicowanie w zakresie objętym egzaminem). Istotną przeszkodą w rozwojowej, pionowej interpretacji wyników polega na tym, że są one oparte na programach różnych przedmiotów i poziomów wiedzy oraz są oceniane za pomocą różnych skal pomiarowych (Tomkowicz, 2008).

Najprostszą skalą pochodną surowych wyników testowania jest **procent opanowanej treści**, oparty na założeniu, że treść kształcenia (danego przedmiotu i szczebla szkoły) dzieli się na sto równych części, a wynik pomiaru pozwala oszacować, ile tych części uczeń *U* opanował. To jawnie antystrukturalne i niekonstruktywistyczne założenie powoduje, że skala procentowa nie cieszy się uznaniem teoretyków pomiaru i jest zaliczana do „pseudonorm wymagań” (Niemierko, 1999, s. 277, 2007, s. 323). Jednak ze względu na prostotę tego założenia i powszechną znajomość rachunku procentowego łatwo przemawia do niespecjalistów, co może skłaniać nas do sięgnięcia po nią na początek.

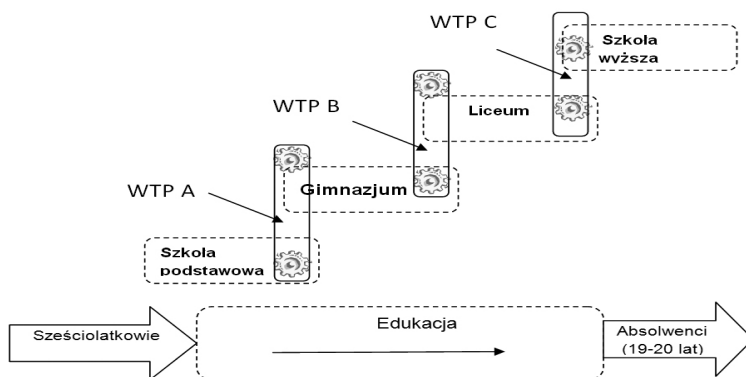
Aneta Wiktorzak (2010) zaproponowała, aby mierzyć rozwój osiągnięć ucznia jako różnicę procentu opanowanej treści na kolejnych szczeblach systemu kształcenia. By zobrazować tę ideę, potraktujemy szkołę jako **wirtualną taśmę produkcyjną** (WTP), która transformuje/przenosi uczniów na wejściu w absolwentów danej szkoły na wyjściu (patrz Walukiewicz, 2010, s. 121). Na rys. 4. pokazano WTP dla przypadku nauczania przedmiotowego (proces P1) w liceum.



Rysunek 4. Wirtualna taśma produkcyjna nauczania przedmiotowego w liceum ogólnokształcącym

Na wejściu WTP na rys. 4. mamy absolwentów gimnazjum z kapitałem ludzkim ocenionym na egzaminie gimnazjalnym. Na wyjściu tej WTP mamy absolwentów liceum, których wartość kapitału ludzkiego została oceniona/oszacowana na maturze. (Analiza zmian wartości kapitału ludzkiego została przedstawiona w rozdziale 6 skryptu Walukiewicza, 2010).

System edukacji można interpretować jako ciąg współpracujących wirtualnych taśm produkcyjnych, jak to przedstawiono na rys. 5.



Rys. 5. Edukacja jako ciąg wirtualnych taśm produkcyjnych

Na wejściu tego ciągu wirtualnych taśm produkcyjnych (WTP A – WTP C) są 6-letnie dzieci rozpoczynające naukę w klasie pierwszej szkoły podstawowej, a na wyjściu są absolwenci, którzy zdali maturę w wieku 19-20 lat. Opis tego ciągu wirtualnych taśm produkcyjnych wraz z wynikami obliczeń zawiera referat Wiktorzak prezentowany na obecnej konferencji PTDE.

Zaprezentowana metoda pozwala na oszacowanie *wzłędnego* przyrostu osiągnięć uczniów, gdyż kolejne szczeble szkoły zawierają nowy materiał kształcenia i *rosnący* poziom wymagań co do podstawowych umiejętności. Skutkiem tego, procentowe wskaźniki opanowania treści kształcenia na ogół *maleją* ze szczebla na szczebel, co kłóci się z intuicyjnym obrazem rozwoju ucznia, ale upodabnia wyniki zastosowania metody pionowego integrowania danych edukacyjnych do wyników wewnątrzszkolnego oceniania osiągnięć uczniów.

6. Czy nauczyciele oceniają kapitał ludzki uczniów?

Skoro kapitał ludzki jest pojęciem tak szerokim, jak przyjęliśmy w tym referacie, a edukacja jest główną drogą jego pozyskiwania, to zasadne staje się pytanie o obecność jego składników w programach kształcenia i w systemach oceniania osiągnięć szkolnych. Tu podejmiemy próbę odpowiedzi tylko na drugą część tego pytania.

Ocena szkolna jest (a) informacją o wyniku uczenia się (stopniem szkolnym) wraz z (b) komentarzem (Niemierko, 2002, s. 184). Obserwacje nauczyciela wybiegające poza wiadomości i umiejętności przedmiotowe ucznia pozaregularnie wpływają na oszacowanie jego osiągnięć (część a oceny), nadając jej charakter **społeczno-wychowawczy** (tamże, s. 212-215), oficjalnie pojawiają się w komentarzu (część b oceny), a najpełniej występują w swobodnej *ocenie opisowej*, często ograniczonej tylko do komentarza, bez stopnia szkolnego (tamże, s. 194-197). Ustny komentarz, aczkolwiek doniosły wychowawczo jako element **oceniania kształtującego**, nie wchodzi w skład **oceniania sumującego** (tamże, s. 186-189), coraz częściej opartego na egzaminie zewnętrznym.

Różnica między egzaminacyjną, jednokryterialną oceną wąskodydaktyczną, a wewnątrz-szkolną wielokryterialną oceną społeczno-wychowawczą powoduje, że ocenianie społeczno-wychowawcze pozwala na równie dobre przewidywanie osiągnięć ucznia w szkole wyższego szczebla, a zapewne i w pracy zawodowej, jak ocenianie wąskodydaktyczne. Ta prawidłowość była wielokrotnie obserwowana w różnych krajach i na różnych szczeblach edukacji (Niemierko, 1969; Young, 2001).

Gdy istniały w Polsce licea pedagogiczne, zgłaszająca się do nich młodzież ze wsi miała znacznie wyższe oceny (społeczno-wychowawcze) na świadectwach ze swoich szkół niż młodzież z miasta, na egzaminie wstępnym uzyskiwała znacznie niższe oceny (wąskodydaktyczne), a już w I klasie odzyskiwała przewagę, która z roku na rok rosła i przenosiła się na praktykę pedagogiczną. Zapewne czynnik motywacyjny („byłam i będę dobrą uczennicą”) podnosił trafność prognostyczną świadectwa.

Podobną sytuację można zaobserwować w Stanach Zjednoczonych. Wyniki najpowszechniej stosowanego Testu Uzdolnień Akademickich (SAT) są skorelowane z osiągnięciami studentów I roku kolegium w wysokości około 0,5, ale wyniki klasyfikacji wewnątrzszkolnej (pozycja ucznia wśród absolwentów szkoły) bywają skorelowane nieco wyżej, około 0,6 (Niemierko, 2002, s. 269). Nie mogą się temu nadziwić specjaliści psychometrii. Jeden z nich (Smith, 2003) tak skomentował to zjawisko:

Badania, jedne po drugich, pokazują, że trafność prognostyczna testów wstępu do kolegium jest w przybliżeniu równa trafności średniej arytmetycznej stopni ze szkoły średniej [...]. Dziwne, że tak jest. Jak znakomite zespoły, złożone z najlepszych na świecie psychometrów, mogą nie móc pokonać niezbornych grup nauczycieli szkół średnich, którzy, w dodatku, mogliby powiedzieć „My wcale nie próbujemy [przewidywać sukcesu w kolegium – BN]”? Odpowiedź na to pytanie mieści się w podobieństwie procesów ustalania ocen w szkole i w kolegium.

„Podobieństwo procesów ustalania ocen”, to jest udział pozapoznawczych składników kapitału ludzkiego w szkole, w kolegium i, zapewne, w pracy zawodowej wyznacza pole rozwoju metodyki oceniania szkolnego. Podjęcie prac nad szacowaniem kapitału ludzkiego i kapitału społecznego uczniów powinno skierować uwagę nauczycieli-wychowawców na następujące kwestie:

1. Uczniowie mają stosowne do wieku wyobrażenie swojej wartości jako kapitału ludzkiego. Ograniczanie oceniania szkolnego do dziedziny poznawczej (wiadomości i umiejętności przedmiotowe), z pomijaniem osiągnięć praktycznych, emocjonalnych i charakterologicznych, lub, co gorsza, niejawne poszerzanie zakresu oceniania o te dziedziny (tamże, r. 14), prowadzi do konfliktu tego wyobrażenia z działalnością szkoły. Docenianie pozapoznawczych dziedzin rozwoju ucznia, choćby jedynie w komentarzu do stopnia, łagodzi ten konflikt.
2. Każdy przedmiot szkolny wiele wnosi do każdej z czterech dziedzin kapitału ludzkiego. Przedmioty humanistyczne (1) usprawniają komunikację międzyludzką, (2) wyposażają w znajomość języków, literatury, historii, stosunków społecznych, (3) uszlachetniają emocje i (4) kształ-

tują charakter i postawy prospołeczne. Przedmioty ściśle (1) usprawniają myślenie teoretyczne i sytuacyjne, (2) poszerzają i pogłębiają znajomość praw natury, (3) ćwiczą ucznia w pokonywaniu trudności i (4) kształtują racjonalny światopogląd. Osiągnięcia ucznia nie sprowadzają się więc do wiadomości i umiejętności przedmiotowych.

3. Każdy przedmiot szkolny wiele wnosi także do przyszłego kapitału społecznego rodziny, zakładu pracy i kręgów działalności społecznej absolwenta: kształtuje umiejętność komunikacji i współdziałania, samokontroli i tolerancji różnicy poglądów i postaw, realistycznej samooceny. Metody pracy grupowej (Niemierko, 2009, r. 8) są jednak wciąż jeszcze zbyt rzadko stosowane w szkole, a sprawdzanie i ocenianie osiągnięć zespołu uczniowskiego są prawdziwym ewenementem (Poręba-Konopczyńska, 2004).
4. Akcentowanie wartości kapitału społecznego *unowocześnia systemy oceniania osiągnięć uczniów*, gdyż jest wybieganiem w dającą się przewidzieć przyszłość ludzkości świata. Obecnie dokonuje się ostra konfrontacja wybujałego **indywidualizmu** kultury euroamerykańskiej z **kolektywistycznym**, opartym na wzajemnej zależności, charakterem społeczeństw azjatyckich. Tę drugą właściwość psychologowie określają jako „zależność od pola”: w spostrzeganiu, reakcjach, sposobie myślenia. Jak stwierdza wnikliwy badacz różnic kulturowych, Richard Nisbett (2010, s. 167), „góra z górą może się jednak zejść. I Wschód, i Zachód, zmierzając ku sobie nawzajem, mogą mieć swój udział w stanowiącym jedną całość świecie, w którym obecne będą – choć przeobrażone – społeczne i poznawcze aspekty obydwu tych regionów [...]”.

Nie wątpię, że wybitni nauczyciele mają pełną świadomość nie tylko bieżących potrzeb swoich uczniów, lecz także nadciągających przemian. Szybki awans krajów i przemysłowych enklaw dalekowschodnich (Japonia, Korea Południowa, Hong-Kong, Szanghaj) w międzynarodowych badaniach osiągnięć uczniów (IEA, OECD/PISA) daje im wiele do myślenia.

7. Od diagnostyki do ewaluatyki

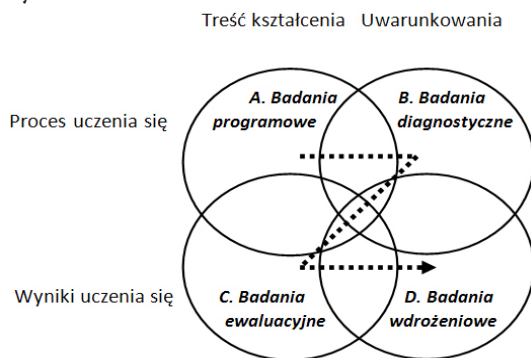
Diagnoza edukacyjna, jako uściślone rozpoznawanie warunków, przebiegu i wyniku uczenia się dokonywane w celu kierowania uczeniem się (Niemierko, 2009, s. 30), jest blisko spokrewniona z **ewaluacją w edukacji**, rozumianą jako zbieranie informacji o warunkach, przebiegu i wynikach działania edukacyjnego dokonywane w celu oceny wartości tego działania i uzasadnienia dotyczących go decyzji (Niemierko, 2009, s. 217). Różnice są następujące:

1. Celem diagnozy edukacyjnej jest wspomaganie procesu kształcenia przez dostarczanie uczniom, ich rodzicom i nauczycielom informacji, której nie byłoby w stanie zdobyć samodzielnie. Celem ewaluacji w edukacji jest dostarczanie właściwym decydentom informacji mogącej posłużyć do skutecznego administrowania pracą podległych im placówek, co obejmuje decyzje organizacyjne, programowe, metodyczne, kadrowe, finansowe i inne decyzje z zakresu polityki oświatowej.
2. Diagnoza jest skoncentrowana na uczeniu się jednostek i grup jednostek, podczas gdy ewaluacja na pierwszy plan wysuwa działanie instytucji edukacyjnych: rodziny, oddziału szkolnego, szkoły, sieci szkół,

- systemu edukacyjnego kraju. Wprawdzie uczenie się pozostaje podstawowym procesem produkcyjnym kapitału ludzkiego w obu wypadkach, ale podejście jednostkowe akcentuje warunki i przebieg uczenia się, a podejście instytucjonalne – jego wyniki.
3. Diagnoza jest dokonywana z reguły własnymi siłami placówki edukacyjnej przy metodologicznym wsparciu ośrodków doradztwa pedagogicznego, a ewaluacja jest zwykle kierowana przez instancje nadrzędne lub instytuty badawcze, przy czym jednak nauczyciele i wychowawcy mogą odegrać główną rolę w dostarczaniu szczegółowej informacji (MacBeath i in., 2003). Tak ułożył się podział czynności rozpoznawczych i decyzyjnych między szczeblami zarządzania edukacją.
 4. Diagnoza ma być „uściślona”, a więc interpretowana pomiarowo, tak dalece, jak to w danym przypadku możliwe, a ewaluacja swobodnie odwołuje się do różnorodnych źródeł informacji: od potocznej przez anegdotyczną i opisową do naukowo-badawczej (Austin i Panos, 1971; Niemierko, 2009, s. 319-320). To poszerza zakres ewaluacji, ale naraża ją na wypaczenia płynące z ogólnosystemowych założeń: autorytarnych, demokratycznych lub liberalnych (Niemierko, 2009, s. 324).

Te różnice powodują, że **diagnostyka edukacyjna**, jako teoria diagnozy edukacyjnej, i teoria ewaluacji, którą moglibyśmy nazwać **ewaluatyką edukacyjną**, rozwijają się równolegle.

Syntezę różnic między diagnostyką edukacyjną a ewaluacją w edukacji przedstawia model czterech rodzajów badań edukacyjnych (Niemierko, 1988, s. 290) umieszczony na rys. 6.



Rys. 6. Uproszczona typologia badań edukacyjnych

Badania diagnostyczne są na rys. 6. zorientowane na *przebieg uczenia się* w określonych warunkach, a badania ewaluacyjne jako zorientowane na *wyniki uczenia się* określonej treści programowej. Cztery typy badań edukacyjnych mocno zachodzą na siebie, co utrudnia ich rozróżnienie, a ułatwia kojarzenie. Układają się w ciąg: od (A) programowych przez (B) diagnostyczne i (C) ewaluacyjne do (D) wdrożeniowych, które czerpią informację z wszystkich poprzednich rodzajów badań.

Najbardziej interesuje nas związek badań ewaluacyjnych z diagnostycznymi. Można sformułować tezę, że wszystkie szerzej znane odmiany **pseudoewaluacji** (Patton, 1981) – tandetna (*quick and dirty*), masywna (*weighty*), towarzyska (*personality-focused*), wydumana (*guesstimate*), przelotna (*eyewash*), wybielająca (*whitewash*), torpedująca (*submarine*), spychająca (*postponement*) i rytualna (*posture*) – są efektem zaniedbania bądź celowego pominięcia diagnozy uczenia się w ocenianych instytucjach na rzecz zbierania danych uzasadniających powzięte z góry, polityczne opinie.

Im pełniejsza jest faza diagnostyczna pedagogicznych badań stosowanych, tym mniejsze jest zagrożenie edukacji ich **efektem zwrotnym** (*backwash effect*), polegającym na wpływie metody ewaluacji osiągnięć uczniów na treść i formę dalszego kształcenia, a w szczególności – na wpływie testów na uczenie się. Szkoły bronią się przed ograniczaniem zakresu diagnozy do wyników kolejnych egzaminów nawet wtedy, gdy te wyniki plasują je na najwyższych pozycjach rangowych. Wyraża to wypowiedź prasowa dyrektora jednej z najlepszych szkół średnich w Polsce, cytowana jako motto naszego referatu.

Wypowiedź dyrektora można interpretować jako upomnienie się o lewą część rys. 1. – uczenie się przez *działanie* i kształtowanie *postaw*. Jego szkoła prowadzi bardzo szeroką działalność intelektualną (wielż z uniwersytetem), językową (międzynarodowa matura, *IB*), krajoznawczą (dyrektor jest geografem) sportową (międzyszkolne ligi) i społeczną (wolontariaty), czym zdobywa ogromną przewagę nad pozostałymi liceami regionu w rankingach prasowych. Odnotowanie i wypunktowanie tych form działalności *szkoły* pozostaje jednak tylko pseudoewaluacją, jeżeli odpowiednie postawy i umiejętności *uczniów* nie są poddane pomiarowi.

Postulat objęcia diagnozą wszystkich czterech składników kapitału ludzkiego (rys. 3.) może zainteresować ewaluatorów, a w miarę dojrzewania odpowiednich procedur także różnego szczebla decydentów edukacyjnych. Na razie mamy do dyspozycji tylko hasła *pomiaru kapitału ludzkiego* i dalej – *pomiaru edukacyjnej wartości dodanej kapitału ludzkiego*, przedstawione w tym referacie do wstępnej dyskusji.

8. Podsumowanie

1. Pojęciom **kapitał ludzki** i **kapitał społeczny** brak jeszcze właściwej reprezentacji pomiarowej, co ogranicza ich użyteczność w edukacji.
2. Punktem wyjścia procedur pomiarowych kapitału ludzkiego i społecznego powinna być **operacjonalizacja** tych pojęć przez dobór obserwowalnych wskaźników.
3. **Najważniejszą właściwością pomiaru zjawisk społecznych jest ich trafność teoretyczna**, w tym **trafność konsekwencyjna** – ustrzeżenie się ujemnych skutków pomiaru.
4. Dziedziny kapitału ludzkiego można wstępnie poklasyfikować na: **fizyczno-praktyczną**, **poznawczo-doświadczałą**, **emocjonalno-motywacyjną** i **charakterologiczno-światopoglądową**.
5. Do pomiaru niematerialnych form kapitału może być zastosowany probabilistyczny model dwuparametrowy **teorii wyniku zadania (IRT)**.

6. Skale pomiarowe kapitału niematerialnego powstaną przez dobór wysoko różnicujących wskaźników w liczbie zapewniającej odpowiednią **rzetelność pomiaru**.
7. **Pionowa i pozioma integracja** danych z egzaminów szkolnych i obserwacji pracy uczniów jest użyteczną metodą diagnozy rozwoju kapitału ludzkiego.
8. Prace nad pomiarem kapitału niematerialnego uczniów spowodują wzrost znaczenia nauczycielskiej **oceny szkolnej**.
9. Postęp metodologiczny diagnostyki edukacyjnej podniesie jej użyteczność w **ewaluacji** pracy placówek oświatowych.
10. Na styku **ekonomii i edukacji** mamy duże rezerwy postępu diagnostyki i ewaluatyki.

Bibliografia:

1. Astin, A.W., Panos, R.J. (1971) *The evaluation of educational programs* [w:] R.L. Thorndike (red.) *Educational measurement. Second edition*. Washington: American Council on Education.
2. Bloom, B. S. (red.) (1956) *Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook I: Cognitive Domain*. New York: McKay.
3. Cronbach, L.J. (1969) *Validation of Educational Measures. Proceedings of the 1969 Invitational Conference on Testing Problems*. Princeton: Educational Testing Service.
4. Dolata, R. (2008) Czy segregacja uczniów ze względu na uprzednie osiągnięcia szkolne zwiększa efektywność nauczania mierzoną metodą EWD? [w:] B. Niemięcki i M.K. Szmigiel (red.) *Uczenie się i egzamin w oczach nauczycieli*. Kraków: PTDE.
5. Doliński, D. (2000) *Emocje, poznanie i zachowanie* [w:] J. Strelau (red.). *Psychologia. Podręcznik akademicki. Tom 2. Psychologia ogólna*. Gdańsk: GWP.
6. Domański, S.R. (1993) *Kapitał ludzki i wzrost gospodarczy*. Warszawa: PWN.
7. Dudzikowa, M. (1993) *Praca młodzieży nad sobą*. Warszawa: Terra.
8. Dudzikowa, M. (2008) *Erozja kapitału społecznego w szkole w kulturze nieufności* [w:] M. Dudzikowa, M. Czerepaniak-Walczak (red.) *Wychowanie. Pojęcia, procesy, konteksty*. Gdańsk: GWP.
9. Gardner, H. (2002) *Inteligencje wielorakie. Teoria w praktyce*. Poznań: Media Rodzina.
10. Glaser, R. (1961) *Instructional Technology and the Measurement of Learning Outcomes*, *American Psychologists*, s. 519–521.
11. Glaser R. (1988) *Przyszłość pomiaru dydaktycznego; Edukacja*, 1.
12. Haertel E.H. (2002) *Standard setting as a participatory process: Implications for Validation of Standard-Based Accountability Programs*, *Educational Measurement: Issues and Practice*, 1.
13. Hambleton, R.K. (1993) *Comparison of Classical Test Theory and Item Response Theory*
14. *and their Applications to Test Development, Educational Measurement: Issues and Practice*, 3.
15. Hornowska, E. (2001) *Testy psychologiczne. Teoria i praktyka*. Warszawa: Scholar.

16. Hulin, C.L., Drasgow, F., Parsons, C.K. (1983/2005). Item Response Theory. Homewood: Dow Jones-Irwin. Tłum. pol. *Wprowadzenie do teorii odpowiedzi na pozycje testu*. [w:] J. Brzeziński (red.) *Trafność i rzetelność testów psychologicznych*. Gdańsk: GWP.
17. Janowski A. (2002) *Poznawanie uczniów. Zdobywanie informacji w pracy wychowawczej*, Warszawa: Fraszka Edukacyjna.
18. Jasińska, A. (2010) *Pomiar gotowości szkolnej uczniów za pomocą skali quasi-obszewacyjnej* [w:] B. Niemierko i M.K. Szmigel (red.) *Teraźniejszość i przyszłość oceniania szkolnego*. Kraków: PTDE.
19. Krathwohl, D.R., Bloom, B.S., Masia, B. (1964) *Taxonomy of Educational Objectives, Handbook II: Affective Domain*. New York: McKay.
20. Kwiatkowski, M. (2005) *Kapitał społeczny* [w:] *Encyklopedia socjologii. Suplement*. Warszawa: Oficyna Naukowa.
21. Lisiecka, Z. (2010) *Badania panelowe w szkołach podstawowych i gimnazjach a rozwój wskaźnika EWD. Doniesienie badawcze* [w:] B. Niemierko i M.K. Szmigel (red.) *Teraźniejszość i przyszłość oceniania szkolnego*. Kraków: PTDE.
22. Larkin J., McDermott J., Simon D.P., Simon H.A. (1980) *Expert and Novice Performance in Solving Physics Problems*, *Science*, s. 1335–1342.
23. MacBeath, J., Schratz, M., Meuret, D., Jakobsen, L. (2003) *Czy nasza szkoła jest dobra?* Warszawa: WSiP.
24. Mager, R.F. (1962) *Preparing Instructional Objectives*. Palo Alto: Fearon.
25. Mager, R.F. (1972) *Goal Analysis*. Belmont: Fearon.
26. Mager, R.F. (1990) *Measuring Instructional Results or Got a Match?* London: Kogan.
27. Marzec-Holka, K. (2009) (red.). *Kapitał społeczny a nierówności – kumulacja i redystrybucja*. Bydgoszcz: Wyd. UKW.
28. Matczuk, A. (2000) *Style poznawcze* [w:] Strelau (red.). *Psychologia. Podręcznik akademicki. Tom 2. Psychologia ogólna*. Gdańsk: GWP.
29. Messick, S. (1980) *Test Validity and the Ethics of Assessment*, *American Psychologist*, 35. Tłum. pol. (2005) *Trafność testu a etyka oceny (diagnozy)* [w:] J. Brzeziński (red.) *Trafność i rzetelność testów psychologicznych. Wybór tekstów*. Gdańsk: GWP.
30. Messick, S. (1989) *Validity*. W: R. L Linn (red.) *Educational Measurement. Third Edition*, Washington: American Council on Education.
31. Muszyński, H. (1971). *Wstęp do metodologii pedagogiki*. Warszawa: PWN.
32. Niemierko B. (1969) *Opowiadaniu nauczyciela w pracy dydaktyczno-wychowawczej*, Warszawa: PZWS.
33. Niemierko B. (1990) *Pomiar sprawdzający w dydaktyce. Teoria i zastosowania*, Warszawa: PWN.
34. Niemierko, B. (1999) *Pomiar wyników kształcenia*. Warszawa: WSiP.
35. Niemierko, B. (2002) *Ocenianie szkolne bez tajemnic*. Warszawa: WSiP.
36. Niemierko, B. (2007) *Kształcenie szkolne. Podręcznik skutecznej dydaktyki*. Warszawa: WAiP.
37. Niemierko, B. (2009) *Diagnostyka edukacyjna. Podręcznik akademicki*. Warszawa: PWN.
38. Nisbett, R.E. (2009) *Geografia myślenia*. Sopot: GWP.
39. Patton, M.Q. (1981) *Creative Evaluation*. London: Sage.
40. Pawłowski, T. (1969) *Metodologiczne zagadnienia humanistyki*. Warszawa: PWN.
41. Popham W.J. (1978) *Setting Performance Standards*. Los Angeles: Instructional Objectives Exchange.
42. Poręba-Konopczyńska, A. (2004) *Test zespołowy z wyposażeniem jako narzędzie oceniania* [w:] B. Niemierko (red.) *Diagnostyka edukacyjna. Teoria i praktyka*. Kraków: PTDE.

43. Schott, F. (1972). Zur Präzisierung von Lehrzielen in Zweidimensionale Aufgabenklassen [w:] K.J. Klauer i in. (red.) *Lehrzielorientierte Tests*. Düsseldorf: Schwann.
44. Smith J.K. (2003) *Reconsidering Reliability in Classroom Assessment and Grading, Educational Measurement: Issues and Practice*, 4.
45. Suchodolski, B. (1947) *Wychowanie dla przyszłości*. Warszawa: PWN.
46. *Standards for Educational and Psychological Testing* (1999). Washington: American Educational Research Association.
47. Szaleniec, H. (red.) (2010) *Teoria wyniku zadania, IRT. Zastosowania w polskim systemie egzaminów zewnętrznych*. Warszawa: Centralna Komisja Egzaminacyjna.
48. Sztompka, P. (2002). *Socjologia. Analiza społeczeństwa*. Kraków: Znak.
49. Tomaszewski, T. (1975). *Człowiek i otoczenie* [w:] T. Tmaszewski *Psychologia*. Warszawa: PWN.
50. Tomkowicz, J. (2008) *Konstrukcja i zastosowanie skal pionowych w pomiarze osiągnięć uczniów* [w:] B. Niemierko i M.K. Szmigiel (red.) *Uczenie się i egzamin w oczach nauczycieli*. Kraków: PTDE.
51. Tyszka, T. (1999) *Psychologiczne pułapki oceniania i podejmowania decyzji*. Gdańsk: GWP.
52. Verhelst, N. (2010) *Probabilistyczna teoria wyniku zadania* [w:] Szaleniec, H. (red.) *Teoria wyniku zadania, IRT. Zastosowania w polskim systemie egzaminów zewnętrznych*. Warszawa: Centralna Komisja Egzaminacyjna.
53. Walukiewicz, S. (2007) *Four Forms of Capital and Proximity, Working Paper WP-3-2007*, Warsaw: Systems Research Institute.
54. Walukiewicz, S. (2010) *Kapitał ludzki. Skrypt akademicki*. Warszawa: IBS PAN.
55. Walukiewicz, S. (w przyg.) *Kapitał społeczny. Skrypt akademicki*. Warszawa: IBS PAN
56. Wendicki, L. (w przyg.) *Kapitał społeczny studenckiej grupy ćwiczeniowej jako czynnik wzrostu kapitału ludzkiego jej członków*. Gdynia: Wyd. AMW.
57. Wiktorzak, A.A. (2009) *Kapitał społeczny szkoły* [w:] B. Niemierko i M.K. Szmigiel (red.) *Badania zagraniczne i wzory międzynarodowe w diagnostyce edukacyjnej*. Kraków: PTDE.
58. Young J.W. (2001) *Differential Validity, Differential Prediction, and College Admissions Testing: A Comprehensive Review and Analysis*. New York: CEEB.