Z badań nad pisemnymi sprawdzianami polonistycznymi

* TADEUSZ PATRZAŁEK
IKNiBO Wrocław

Sprawdzanie osiągnięć szkolnych, funkcjonujących jako element systemu kształcenia i wychowania, oddziałuje na ten system swoiście wzmacniająco; przejawia się to m.in. w dostosowywaniu uczenia do treści i form sprawdzania. W skrajnej postaci może to nawet prowadzić do postawy, którą wyraża takie oto zdanie ucznia: „Uczę się tego i tak, czego i jak wymaga ode mnie nauczyciel". Sprawdzanie osiągnięć ma zatem szczególnie mocne powiązania zwrotne z pozostałymi elementami systemu, a doskonalenie sprawdzania może skutecznie wpływać na ulepszanie całego procesu kształcenia i wychowania.

W artykule omawiam niektóre wyniki badań nad jednym ze sposobów sprawdzania osiągnięć polonistycznych, w tym głównie literackich, choć nie oddzielam ich od językowych, aby możliwe było ich porównanie i aby zarysował się obraz całości osiągnięć polonistycznych wymaganych w pisemnych sprawdzianach.

1. PRZEDMIOT, CELE, CHARAKTER I ZAKRES BADAŃ

Wiadomo z potocznej obserwacji, że w praktyce szkolnej funkcjonują dziś dwie podstawowe formy pisemnego sprawdzania osiągnięć uczniów z języka polskiego — wypracowanie i sprawdzian, nazywany też „kartkówką".

Wypracowaniem nazywamy tutaj dłuższą, wielofunkcyjną wypowiedź pisemną na określony temat, której celem jest m.in. sprawdzenie umiejętności wytwarzania tejże wypowiedzi, a ocenie szkolnej podlega sam wytwór pisania. Wypracowanie ma długą tradycję w szkole oraz bogate piśmiennictwo metodyczne i w zasadzie pozostaje poza przedmiotem zainteresowania w niniejszych badaniach.

Przyjmuję, że *sprawdzian pisemny* jest odmianą *testu osiągnięć szkolnych*, tzn. zbiorem stosunkowo luźno ze sobą powiązanych zadań, tak dostosowanych do treści nauczania, że na podstawie wyników testowania można ustalić, w jakim stopniu te treści są opanowane przez

uczniów. Sam test jako wytwór (konstrukt), czyli forma wypowiedzi ucznia, nie podlega ocenie szkolnej.

Badania nad polonistycznymi sprawdzianami zostały tutaj podjęte po raz pierwszy w Polsce, a ich ogólnym celem jest opisanie wybranych właściwości sprawdzianów jako testów osiągnięć szkolnych, tzn tych właściwości, które przede wszystkim warunkują możliwość wypełniania przez sprawdziany wyznaczonej im funkcji dydaktycznej, wynikającej z definicji testu osiągnięć szkolnych.

W szczególności zaś celem badań było wstępne, orientacyjne ustalenie rozmiarów zjawiska stosowania sprawdzianów polonistycznych oraz próba odpowiedzi na dwa podstawowe pytania:

— jakie są konstrukcyjne właściwości sprawdzianów w porównaniu z wymogami współczesnego testu osiągnięć szkolnych?
— jakie osiągnięcia sprawdza ta odmiana testu w świetle celów nauczania formułowanych przez współczesną dydaktykę języka polskiego?

Celem praktycznym, wynikającym z odpowiedzi na te dwa podstawowe pytania, jest sformułowanie potrzeb w zakresie doskonalenia pisemnego sprawdzianu jako testu polonistycznego. Tak wyznaczony cel praktyczny powoduje z kolei, że analiza właściwości sprawdzianów odkrywa przede wszystkim ich wady i braki.

Przeprowadzone badania miały charakter empiryczny i dzieliły się na dwie fazy: 1. zbierania materiałów, w postaci polonistycznych sprawdzianów nauczycielskich, oraz danych dotyczących ich stosowania w wybranych szkołach; 2. ilościowego opracowania zebranych danych i jakościowych analiz.

Badania ze względów technicznych i treściowych zostały ograniczone do sprawdzianów dla szkół średnich. Sprawdziany dla klas młodszych trzeba by potraktować oddzielnie, stanowią one bowiem, jak się zdaje, inne w sensie jakościowym i ilościowym zjawisko oraz wymagałyby nieco odmiennych procedur i technik badawczych. Stosowanie sprawdzianu w szkole średniej uznaje się tutaj za charakterystyczny przykład posługiwania się testem we współczesnym, polskim szkolnictwie.

Badania wprowadzające, dotyczące rozmiarów zjawiska, są oparte na materiale zebranym w liceach ogólnokształcących Dolnego Śląska; w badaniach podstawowych, opisujących właściwości sprawdzianów, zostały wykorzystane materiały zebrane w całym kraju.

2. STOPNIE ZA SPRAWDZIANY W OCENIE UCZNIÓW

Aby oszacować rozmiary i istotność zjawiska, jakim jest posługiwanie się sprawdzainami na lekcjach języka polskiego, przeprowadzi-

łem badania wstępne w celowej próbie 94 klas liceów ogólnokształcących. W klasach tych wystawiono ogółem ok. 46000 stopni, w tym za sprawdziany — ok. 9700, co stanowi 21% wszystkich wystawionych stopni. Oznacza to, że w badanej grupie średnio co piąty stopień w dzienniku był wystawiony przy zastosowaniu sprawdzianu.

Udało się też sformułować następujące wnioski dotyczące zaobserwowanych korelacji pomiędzy zmiennymi składającymi się na styl pracy nauczyciela a liczbą stosowanych przez niego sprawdzianów:

1. Nie widać wyraźniej współzależności pomiędzy stosowaniem sprawdzianów a poziomem pracy nauczyciela, wyrażającym się w wystawianej mu ocenie podczas hospitacji i wizytacji.
2. Stosunkowo najczęściej posługują się sprawdzianami nauczyciele o średnim stażu pracy (8—16 lat), najrzadziej zaś — początkujący i z bardzo długim stażem.
3. Można dostrzec pewną dodatnią korelację pomiędzy stopniem zorganizowania warsztatu polonisty a korzystaniem ze sprawdzianów.
4. Najwyższą, dodatnią korelację wykazuje stosowanie sprawdzianów z oceną pracowni polonistycznej, tzn. że im lepiej jest wyposażona pracownia, tym na ogół częściej jej opiekun korzysta ze sprawdzianu.

Okazało się jednak, że stopnie wystawione w dzienniku mają dla nauczyciela różną wartość w tym sensie, że np. „trójka" wystawiona w czasie odpytywania bieżącego nie ma tej samej wartości co „trójka" za klasówkę. Aby zbadać, jak poloniści wartościują stopnie za sprawdziany wśród pozostałych stopni, przygotowałem ankietę proponującą nauczycielom skonstruowanie *porządkowej skali wartościowania* z wyróżnionych przez nich grup stopni wystawianych w dziennikach lekcyjnych. (Trzeba tu wyjaśnić, że usiłowania zmierzające do skonstruowania skali *przedziałowej* nie powiodły się, albowiem ankietowanym trudno było się zdecydować na jakiekolwiek ważenie stopni należących do poszczególnych grup, tzn. nie potrafili określić, o ile czy ile razy stopień z danej grupy jest w ich ocenie ważniejszy od stopnia z innej grupy.)

Zamieszczona niżej tabela 1 zawiera dane z 43 ankiet. Liczby w kolumnach oznaczają, ile razy w danym miejscu skali wymieniali nauczyciele poszczególne grupy stopni.

Pierwsze, najwyżej wartościowane miejsce wypracowania klasowego jest bezdyskusyjne tabela 1). Sprawdzian rzadko był wymieniany na pierwszym miejscu, ale też ani razu nie plasował się poniżej czwartego miejsca na skali. Zastanawia wysoka ranga stopni za aktywność, w których zawierają się wyniki bieżącej obserwacji uczniów. W uwagach do ankiety nauczyciele wspominali, że ranga obserwacji byłaby jeszcze

Tabela 1

Wartościowanie stopni przez nauczycieli

Lp.	Stopnie za:	Pozycja na skali					
		I	II	III	IV	V	VI
1	Wypracowanie klasowe	33	6	4	—	—	—
2	Sprawdzian (test)	2	14	10	17	—	—
3	Aktywność	8	—	14	8	10	3
4	Odpytywanie	1	12	10	12	8	—
5	Zeszyt (z pracami domowymi)	—	3	2	4	11	3
6	Recytacja	—	—	—	1	2	5
7	Współczesne życie kulturalne (uczestnictwo)	—	—	1	1	2	—
8	Prace nadobowiązkowe	—	—	2	—	1	—

wyższa, gdyby można ją wprowadzić w sposób bardziej zorganizowany i mniej subiektywny. W tychże uwagach ankietowani zastrzegali się, jakby nie ufając swoim różnego rodzaju „kartkówkom", że jeśli wysoko ocenią stopnie za sprawdziany, to pod warunkiem, że sprawdziany (testy) są właściwie opracowane. Inne, nie wymienione wprost w ankiecie grupy stopni (w tabeli 1 pod linią przerywaną) nie odegrały w konstruowanej skali wartościowania istotniejszej roli.

Jakkolwiek by nie patrzeć na uzyskane dane, można chyba stwierdzić, że ranga stopni za sprawdziany (testy) w opinii nauczycieli-polonistów jest wysoka; po stopniach za wypracowanie klasowe plasują się one wśród trzech najważniejszych, zdaniem nauczycieli, grup stopni. Wniosek ten, łącznie z poczynionymi wcześniej spostrzeżeniami dotyczącymi liczby stopni za sprawdziany, upoważnia do uogólnienia, że sprawdzianowa (testowa) forma kontroli wyników nauczania ma już dziś istotny wpływ na ocenianie uczniów z języka polskiego; niektóre inne spostrzeżenia nasuwają zaś przypuszczenie, że ten wpływ będzie się jeszcze zwiększał w przyszłości. Warto zatem przyjrzeć się jakości sprawdzianów, które tak istotnie współdecydują o ocenach uczniów.

3. SPRAWDZIANY PUBLIKOWANE JAKO PODSTAWOWY PRZEDMIOT BADAŃ

Dla zbadania jakości stosowanych przez polonistów sprawdzianów należało wpierw zgromadzić ich odpowiednią liczbę. Powinna to być jednocześnie próba reprezentatywna dla wszystkich sprawdzianów stosowanych w określonym czasie i na określonym terenie badań. Postępowanie takie uniemożliwia zrozumiała niechęć nauczycieli do udostępniania komukolwiek z zewnątrz materiałów przeznaczonych wyłącznie do ich własnego użytku. W dodatku nie byłoby nigdy wiadomo, jaki

procent wszystkich skonstruowanych i wykorzystanych sprawdzianów stanowiłyby zebrane, mimo trudności, sprawdziany nauczycielskie.

Dlatego zdecydowałem się uczynić przedmiotem badań sprawdziany w jakiś sposób *publikowane*, np. powielane jako materiały pomocnicze dla nauczycieli, drukowane w zbiorach typu „z moich doświadczeń", zamieszczane w czasopiśmie przedmiotowym.

Sprawdziany publikowane nie mieszczą się w stosowanej dotąd w literaturze przedmiotu typologii testów, zajmują miejsce pośrednie pomiędzy testem *standaryzowanym* a *nieformalnym*[1], są nauczycielskimi testami szerokiego użytku. Przyjąłem założenie, że mogą one być reprezentatywnym materiałem do badań nad jakością wszystkich sprawdzianów nauczycielskich, a w każdym razie — że nie są one jakościowo gorsze pod względem podstawowych wymogów stawianych dziś przez dydaktykę tego rodzaju narzędziom pomiaru od sprawdzianów nie publikowanych.

Przedmiotem badań stały się sprawdziany publikowane w Polsce w latach 1971—1795[2]. Za przyjęciem tych dat, ograniczających materiał w czasie, przemawiał fakt, że właśnie w roku 1971 wszedł w życie nowy program nauczania, wprowadzający po raz pierwszy — jako obowiązkową — formę sprawdzianową kontroli wyników nauczania. W roku 1975 zaś ukazały się w Polsce pierwsze znaczące publikacje z zakresu testów osiągnięć szkolnych, które zapoznawały z dorobkiem zagranicznym w tym zakresie i prawdopodobnie zaczną istotnie oddziaływać na sprawdziany nauczycielskie. Na rok 1976 przypada z kolei wydanie pierwszych polonistycznych testów standaryzowanych.

Dzięki takiemu ograniczeniu w czasie można było uznać opisywane sprawdziany za zjawisko rodzime, wyrosłe z zapotrzebowania praktyki szkolnej i pod jej wpływem ukształtowane. Należy się spodziewać, że ich analiza dostarczy danych nie tylko o stosowanych narzędziach kontroli wyników, ale także — o treściach i poziomie nauczania przedmiotu na określonym etapie funkcjonowania polskiej szkoły.

Ostatecznie przedmiotem badań stało się 18 sprawdzianów złożonych z 542 zadań z zakresu nauczania języka (które dalej będę oznaczał skrótem „J") oraz 49 sprawdzianów złożonych z 870 zadań z zakresu literatury (oznaczonych dalej skrótem „L"). Podane liczby zadań nie są dokładnie zgodne z liczbami formalnie wydzielonych zadań w analizowanych sprawdzianach, albowiem zachodziła czasami konieczność skorygowania oczywistych pomyłek redakcyjnych w publikacjach.

---

[1] Terminów tych i wielu następnych, dotyczących pomiaru dydaktycznego, używam na ogół w takim znaczeniu, w jakim występują one w książce B. Niemierki, *Testy osiągnięć szkolnych*. Warszawa 1975.

[2] Nie były brane pod uwagę sprawdziany publikowane w tych latach, ale oparte na starym programie nauczania sprzed roku 1971.

Ze względów technicznych odrzucane też były pojedyncze zadania z języka, bez żadnego uzasadnienia wtrącane w sprawdziany z literatury.

W sumie zatem badania objęły 67 sprawdzianów publikowanych, złożonych z 1412 zadań, stanowiących mniej więcej całość tego rodzaju publikacji, które ukazały się w Polsce w latach 19771—19775. Zebrany materiał stanowi więc *populację*, a nie próbę, co uwalnia od statystycznych procedur zmierzających do nadania próbie reprezentatywności i pozwala traktować badania jako *wyczerpujące*.

4 SUMUJĄCY CHARAKTER SPRAWDZIANÓW PUBLIKOWANYCH

Jak już powiedziano, sprawdziany publikowane są odmianą testu osiągnięć szkolnych, a te, wg przyjętej tutaj typologii, dzielą się na różnicujące i sprawdzające[3]. Nie ma powodu dopatrywać się w którymś spośród 67 sprawdzianów publikowanych cech testu różnicującego; nie upoważniają do tego ani odautorskie charakterystyki, ani właściwości konstrukcyjne samych sprawdzianów celowo konstruowanych dla celów selekcyjnych. Dlatego traktuję je dalej wszystkie jako testy sprawdzające, co jest na ogół zgodne z ich odautorskimi określeniami uwidocznionymi w tytułach.

Tytułom tym zresztą warto przyjrzeć się bliżej, wyrażają one bowiem pewne charakterystyczne tendencje, potwierdzające się w analizie treści i formy sprawdzianów.

We wszystkich tytułach występuje ogólne określenie rodzaju narzędzia („sprawdzian" lub „test"), z rzadka tylko łączone z nazwaniem zakresu bądź charakteru treści zawartych w sprawdzianie. Tylko bowiem w dwunastu (4-J i 8-L) tytułach występują próby ich materiałowego ukonkretnienia, np. „Sprawdzian wiadomości z gramatyki opisowej *(składnia)*" lub „Test badający znajomość *utworów Wł. Broniewskiego z okresu 20-lecia*". W tytułach pozostałych 55 sprawdzianów określenie zakresu materiału jest bardzo ogólnikowe, np.: „Sprawdzian z *gramatyki*", a najczęściej wskazuje tylko poziom zastosowania, np. „Sprawdzian *dla II klasy licealnej*". Tytuły zatem nie precyzują tematów ani nie określają wyraźnie zakresu materiału sprawdzianów, umożliwiając pomieszczanie w nich bardzo różnorodnych treści.

Tylko 10 razy posłużyli się autorzy w tytułach nazwą „test". Trudno znaleźć jakąkolwiek inną prawidłowość w stosowaniu nazw „sprawdzian" i „test" nad tę, iż drugą nazwę łączy się z zastosowaniem zadań zamkniętych wyboru wielokrotnego (co potwierdza przeprowadzona da-

[3] Por. B. Niemierko, op. cit., s. 27.

lej analiza form zadań), i w ogóle — z narzędziem bardziej zaawansowany mkonstrukcyjnie.

W tytułach sprawdzianów stosuje się też określenia natury taksonomicznej, przy czym prawie zawsze, bo w 38 na 40 przypadków, konstruktorzy używają nazwy „wiadomości". Trudno przypuszczać, aby celowo chcieli ograniczyć swoje konstrukty do poziomu taksonomicznego wiadomości, tym bardziej że w sprawdzianach tych zdarzają się także zadania z wyższego poziomu umiejętności. Dlatego posługiwanie się terminem „wiadomości" trzeba chyba uznać za nieświadome, odziedziczone z tradycji językowej[4].

Ze względu na przeznaczenie i wielkość zakresu materiału można podzielić testy na: a) diagnostyczne, b) sumujące, c) mieszane (pośrednie między diagnostycznymi i sumującymi). Testy *diagnostyczne*, mające węższy zakres materiału (od kilku do kilkunastu lekcji), stosuje się do sprawdzania procesu kształcenia przed zakończeniem danej fazy, co umożliwia jego modyfikację. Testy *sumujące*, mające szerszy zakres materiału (obejmują duży dział programu, okres, klasę), stosuje się do sprawdzenia sumującego na zakończenie danej fazy procesu kształcenia, celem kontroli jego wyników. W testach diagnostycznych zadania pokrywają materiał na tyle gęstą siatką, że na podstawie wyników analizy testowania można badać przyczyny niepowodzeń uczniów, wespół z analizą warunków, w jakich przebiegał proces kształcenia. Zadania w tekście sumującym reprezentują zwykle w pojedynkę zbyt duże obszary treści programowych, aby za ich pomocą wystarczająco dokładnie można było lokalizować przyczyny sukcesów i porażek dydaktycznych. Wyniki sprawdzania sumującego dają tylko ogólną orientację w poziomie wiedzy uczniów.

W tabeli 2 podzielono sprawdziany podług zakresu materiału i możliwości wykorzystania uzyskanych za ich pomocą wyników testowania.

Tabela 2

Podział sprawdzianów

Sprawdziany	J	L	Razem
A. diagnostyczne	1	6	7
B. sumujące	13	41	54
C. mieszane	4	2	6

Powyższa tabela potwierdza, zaobserwowaną już w tytułach, tendencję polonistów do konstruowania sprawdzianów sumujących; do ta-

[4] Myślę tu o dydaktycznej tradycji językowej, wywodzącej się u nas m.in. z książek M. Grzywak-Kaczyńskiej i J. Pietera.

kiego rozszerzania zakresu ich treści i „rozrzedzania" w ten sposób pokrywającej ją siatki zadań, że wyników sprawdzania takim testem nie można wykorzystać dla celów kształcących.

Na podstawie tej właściwości można przypuszczać, iż są one wykorzystywane nie do kontrolowania prawidłowości procesu kształcenia, lecz tylko jako usprawnienie przy wystawianiu stopni, a w konsekwencji — dla ułatwienia i ewentualnie doskonalenia przebiegu oceniania. (Przypuszczenie to potwierdza fakt, że w dziennikach lekcyjnych w zasadzie brak śladów wprowadzenia na forum klasy analiz wyników testowania, które byłyby odpowiednikami np. powszechnie praktykowanego poprawiania wypracowań klasowych. Analiza wyników testowania, jako jednostka metodyczna, nie funkcjonuje w dotychczasowej praktyce szkolnej.)

5. FORMY ZADAŃ

W toku historycznego rozwoju ukształtowały się dziesiątki różnych form zadań testowych. Żeby zbytnio nie uszczegółowiać i nie komplikować ich typologii, wyróżniam — za literaturą przedmiotu — tylko sześć zasadniczych form. Tabela 3 prezentuje, które z nich i ile razy wykorzystywali konstruktorzy polonistycznych sprawdzianów publikowanych.

Tabela 3

Formy zadań testowych

Zadania	J	L	Razem
A. otwarte (O):			
1. krótka odpowiedź (KO)	447	683	1130
2. z luką (L)	93	106	199
3. rozprawka (R)	—	33	33
B. zamknięte (Z):			
4. wyboru wielokrotnego (WW)	2	41	43
5. na dobieranie, porządkowanie (D)	—	7	7
6. prawda — fałsz (PF)	—	—	—

Tabela 3 uwidacznia prawie zupełny brak zadań zamkniętych (ledwie 3,5%). Trudno to wyjaśnić jedynie faktem, że są one mniej opłacalne przy testowaniu małych grup uczniów, skoro zostały opublikowane, oraz tym, że trudniej je konstruować; choć rzeczywiście — częściej pojawiają się one w sprawdzianach bardziej zaawansowanych, nazywanych przez autorów „testami". Wydaje się, że niepopularność zadań zamkniętych w sprawdzianach polonistycznych ma swoje głębsze przyczyny, m.in. w specyficznych właściwościach treści przedmiotu,

a także — w uprzedzeniu polonistów do tej formy zadań. Jest to jednak tylko hipoteza, którą trzeba by sprawdzić, wykorzystując inny rodzaj materiału i stosując inne metody badań.

Jak widać, konstruktorzy prawie w ogóle nie posługują się zadaniem otwartym w formie rozprawki. Trzeba bowiem dodać, że część spośród 33 zadań (tylko z literatury!) zakwalifikowanych jako rozprawki, „stała się" rozprawkami wbrew zamierzeniom konstruktorów, z powodu nadmiernego rozszerzenia zakresu treści zadań krótkiej odpowiedzi, a więc — na skutek błędu konstrukcyjnego. Poloniści stosują w swej praktyce albo wypracowania, albo testowe zadania krótkiej odpowiedzi, unikają natomiast formy pośredniej, jaką jest właśnie zadanie testowe w formie rozprawki.

Dominuje zatem w sprawdzianach publikowanych zadanie krótkiej odpowiedzi, stosunkowo łatwe w konstruowaniu, ale też najbardziej naturalne, zbliżone w formie do pytań, jakie nauczyciele stosują na lekcji.

Zadania mają liczne wady i błędy konstrukcyjne (około dwudziestu różnych typów usterek), ale niczym specyficznym się one nie wyróżniają; dają się pomieścić w opisywanej przez literaturę przedmiotu typologii błędów zadań testowych[5].

6. TECHNICZNE WŁAŚCIWOŚCI SPRAWDZIANÓW PUBLIKOWANYCH

Żaden z analizowanych sprawdzianów nie był publikowany w dwu wersjach (wariantach). W zasadzie brak im też charakterystyk (przewodników), które opisywałyby ich indywidualne właściwości i określały zakres zastosowania (do czego się nadają). Sprawdziany publikowane w zbiorach miewają wprawdzie odautorskie czy redakcyjne wstępy, ale najczęściej powtarzającą się w nich myślą jest usprawiedliwianie publikacji za jej niedostatki. Zamiast ujednoliconych, opisywanych w charakterystykach zasad punktowania odpowiedzi, autorzy najczęściej objaśniają sposób punktowania oddzielnie przy każdym zadaniu, na marginesach egzemplarzy sprawdzianów. W ogóle opracowanie redakcyjne sprawdzianów jest mało ekonomiczne; np. zadania są często tak zredagowane, że nie przewidują wykorzystania „kart odpowiedzi", lecz wypełnianie luk (wykropkowań) pozostawionych na egzemplarzu. Nauczyciel, chcąc powielić sprawdzian i zachować egzemplarze do wielokrotnego użytku, musiałby dokonywać znacznych zmian redakcyjnych w stosunku do tekstu publikowanego.

Wszystkie sprawdziany publikowane zaopatrzone zostały w *normy testowe*, czyli zasady przeliczenia punktów na stopnie. Są to wszędzie

[5] Błędy te opisuję szczegółowo w książce pt. *Metodyka testu polonistycznego*. Warszawa 1977.

normy automatyczne i zwyczajowo w naszym kraju przyporządkowane sprawdzianom, tzn. na stopień dostateczny wymaga się ok. 50% punktów możliwych do zdobycia, niezależnie od charakteru i trudności sprawdzianu.

Nieoczekiwanie okazało się, że konstruktorzy-poloniści mają duże trudności językowe, głównie — w zwięzłym i jednoznacznym formułowaniu poleceń.

7. WŁAŚCIWOŚCI TAKSONOMICZNE ZADAŃ

Podstawową, wyróżnianą w teorii i praktyce właściwością treści zadania jest jego wartość *taksonomiczna*. I znów, aby zbytnio nie rozdrabniać podziałów, w zestawieniu liczbowym (tabela 4) ograniczam się do ukazania przynależności treści zadań do kategorii i poziomów taksonomicznych, nie zajmując się ich szczegółową klasyfikacją. Ponieważ kwalifikowanie taksonomiczne treści zadań na obecnym etapie rozwoju dydaktyki jest procedurą po części subiektywną, nie posługuję się dokładnymi liczbami, lecz zaokrąglonymi wskaźnikami procentowymi zadań, którym przypisuję określone własności taksonomiczne. Upraszczam też schemat taksonomii do postaci umożliwiającej jednoczesne pomieszczanie w nim treści zadań z obu działów przedmiotu, tzn. języka i literatury.

Tabela 4

Kwalifikacja taksonomiczna treści zadań (w procentach)

Kategorie i poziomy	J	L	Razem
I. znajomość (wiadomości)	50	71	60,5
II. umiejętności, w tym na poziomie:	50	26	38
— niższym (sprawności)	37	23	30
— wyższym (problemy)	13	3	8
III. postawy	0	3	1,5

W sprawdzianach publikowanych w zasadzie nie występują zadania dotyczące postaw, co nie powinno zaskakiwać, jeśli się zważy, iż w ogóle testy osiągnięć szkolnych, do których zaliczyliśmy sprawdziany, mało nadają się, w ich dzisiejszej postaci, do pomiaru osiągnięć w tej najwyższej kategorii.

Góruje nad innymi kategoriami znajomość treści zadań (wiadomości) głównie w dziale literatury. Jest to często znajomość najniżej w taksometrii wartościowana, tzn. proste pamiętanie wiadomości historycznoliterackich i kontekstowych w ogóle (co w uproszczonym schemacie zestawienia nie mogło być uwidocznione).

W idealnej proporcji układają się zadania należące do kategorii wiadomości i umiejętności w dziale języka. Nie zawsze są to jednak umiejętności, jak się dalej okaże, na których dziś najbardziej powinno polo-niście zależeć.

Wewnątrz kategorii umiejętności, szczególnie w dziale literatury, zdecydowanie przeważają liczebnie zadania z poziomu niższego. Oznacza to, że autorzy sprawdzianów dużo częściej pytają na przykład o gramatyczną poprawność tekstu lub o zgodny z normami poetyki opis utworu niż o właściwości stylistyczne wypowiedzi czy interpretację dzieła.

Ogólnie trzeba zauważyć, że lepszą strukturę pod względem taksonomicznych właściwości treści mają sprawdziany językowe niż literackie.

8. ZASADNICZE TYPY MATERIAŁU PROGRAMOWEGO W ZADANIACH

Kategoria taksonomiczna zadania jest rezultatem nie tylko zamierzonego działania konstruktora, ale jest też uzależniona od jakości samego materiału, a także od ilości włożonej w zadanie pracy i od możliwości konstruktora. Wiadomo, że na ogół tym trudniej ułożyć zadanie testowe, im wyższą kategorię taksonomiczną ma ono reprezentować. Trzeba zatem przyjąć, że na opisany wyżej rozkład zadań w kategoriach taksonomicznych mogły mieć wpływ trudności „warsztatowe", wynikające z niewystarczających kwalifikacji autorów sprawdzianów.

Dlatego spróbowałem zbadać, na jakie typy materiału zostały ukierunkowane zadania. Aby postępować właściwie w tym zakresie, nie trzeba być wytrawnym konstruktorem testu, wystarczy po prostu respektować zasadnicze cele programu nauczania i aktualne tendencje w dydaktyce.

Przyjąłem, że cały materiał programowy poświęcony językowi jest ukierunkowany albo na cele *normatywne* — uczenie praktycznego posługiwania się językiem, jego poprawności i sprawności, albo na *opisywanie* języka — a więc na wiedzę o języku. Zadania w sprawdzianach publikowanych reprezentują te dwa typy materiału następująco:

— materiał o charakterze normatywnym — 74 (12%) zadań,
— materiał o charakterze opisującym — 468 (88%) zadań.

Widać z tego, że w sprawdzianach językowych dominuje wiedza teoretyczna, odkrywająca strukturę języka (głównie gramatyka, która wypełnia w ponad 80% sprawdziany językowe), nad wiedzą praktyczną, uczącą mówić i pisać. Takie ustawienie kierunku sprawdzania razi nie tyle w zestawieniu z dotychczasowym programem nauczania, co z najnowszymi tendencjami w dydaktyce przedmiotu, których odzwiercie-

dleniem jest nowy, projektowany program szkoły dziesięcioletniej, zdecydowanie wysuwający na plan pierwszy kształcenie języka.

Materiał działu literackiego również podzieliłem dychotomicznie: na przynależący do zakresu *utworu* i do zakresu najszerzej rozumianego *kontekstu*. Zadania sprawdzianów publikowanych rozłożyły się następująco w obu grupach:

— o utworze — 514 (59%),
— o kontekście — 356 (41%).

Rozkład ten należałoby ocenić pozytywnie, bowiem problematyka utworu (kosztem kontekstu) preferowana jest przez program, naukę o literaturze i dydaktykę przedmiotu. W tej sytuacji udział wiedzy spoza utworu i tak wydaje się za duży. Gorzej wygląda sprawa, gdy wewnątrz materiału o utworze wyróżnimy to, co dotyczy jego *literackości*. Wówczas okazuje się, że ten najbardziej specyficzny dla literatury materiał uwzględnia zaledwie ok. 10% wszystkich zadań. Dominują natomiast zadania dotyczące świata przedstawionego (33%). Znaczny jest również udział charakterystycznych dla sprawdzianów zadań polecających rozpoznanie utworu na podstawie cytatu, postaci, zdarzenia (ponad 10%); niby pytają one o utwór, ale mają małą wartość dydaktyczną.

Wśród zagadnień o kontekście z kolei znaczną ilościowo pozycję zajmują zadania pytające o tzw. życie kulturalne na tle historycznym (ok. 17%), genologię (11%) i biografię (6%).

Dzieląc tradycyjnie materiał o utworze na tyczący formy i tyczący treści, zauważa się w sprawozdaniach publikowanych zdecydowaną przewagę tego drugiego.

9. SPRAWDZIANY PUBLIKOWANE I NIE PUBLIKOWANE

W końcowej fazie badań spróbowałem sprawdzić zasadność przyjętego na początku założenia, iż sprawdziany publikowane reprezentują trafnie wszystkie sprawdziany nauczycielskie i że nie są one jakościowo gorsze od sprawdzianów nie publikowanych, jakimi na co dzień posługuje się nauczyciel. W tym celu wylosowałem spośród zgromadzonych sprawdzianów nie publikowanych 9 językowych i 12 literackich, co stanowi odpowiednio 50% i 25% sprawdzianów publikowanych i poddałem je analizie.

Okazało się, że wszystkie podstawowe właściwości sprawdzianów nie publikowanych są podobne, a wskaźniki liczbowe bliskie lub wyraźnie gorsze. I tak np. zadania czysto wiadomościowe stanowią w nich 72% (J) i 83% (L), w sprawdzianach publikowanych stanowiły one odpowiednio 50% i 71%. Umiejętności z poziomu wyższego wypełniają 5% (J) i 0,4%

(L) zadań w sprawdzianach nie publikowanych, a 13% (J) i 3% (L) w publikowanych. Zupełnie nie funkcjonuje w sprawdzianach nie publikowanych kategoria postaw, nawet w dziale literatury.

Gdy zadania sprawdzianów publikowanych były rozproszone po całym zakresie treści programowych i bardziej równomiernie, choć w pojedynkę je reprezentowały, to zadania w sprawozdaniach nie publikowanych występują często jakby wiązkami, gęściej pokrywając jedne, a zupełnie opuszczając inne partie programu. Natomiast sam materiał w obu rodzajach sprawdzianów jest jednakowo obszerny.

Zadania zamknięte wyboru wielokrotnego pojawiły się tylko u tych autorów, o których skądinąd było wiadomo, że mają dostęp do powielaczy lub dysponują egzaminatorami, zakupionymi albo skonstruowanymi w pracowniach własnym przemysłem.

W zasadzie wszystkie sprawdziany nie publikowane mają dwie wersje, zróżnicowane z reguły najprościej, tzn. przez odwrócenie kolejności lub nieznaczne zmiany redakcyjne zadań. (Można chyba przypuszczać, że sprawdziany publikowane konstruowano również w dwu wersjach, a były tylko drukowane w jednej wersji ze względów oszczędnościowych.)

10. WNIOSKI

Sprawdziany publikowane i nie publikowane stanowią zatem tę samą jakość i można uznać, iż niżej sformułowane wnioski zostały wyciągnięte z analizy 88 pisemnych sprawdzianów polonistycznych, złożonych w sumie z 1834 zadań.

a) Pisemny sprawdzian polonistyczny, choć ukształtowany poza teorią testu, w naszych rodzimych warunkach nie stanowi jakiejś nowej, specyficznej jakości; jest mniej zaawansowaną konstrukcyjnie odmianą testu osiągnięć szkolnych, w zasadzie złożoną z zadań otwartych krótkiej odpowiedzi, przeznaczoną do sprawdzania sumującego
b) Badania ujawniły w nim tyle istotnych braków, zarówno z punktu widzenia wymogów testu, jak i treści przedmiotowych, iż stają pod znakiem zapytania korzyści płynące z jego stosowania. Ponieważ jednak sprawdzian jest już dziś istotnym elementem procesu dydaktycznego, a jego rola zdaje się jeszcze wzrastać, należy uruchomić wszelkie środki dla jego doskonalenia.
c) Stan obecny sprawdzianów i duże wymagania, jakie stawia współczesna dydaktyka przed konstruktorem, zdają się wskazywać, że z powodów merytorycznych i technicznych nauczyciel nigdy nie będzie mógł zaspokoić we własnym zakresie zapotrzebowań na dobre

testy osiągnięć szkolnych. Dlatego wydaje się, że powinien on je otrzymywać gotowe, i to w wydaniu „listkowym", dla każdego ucznia, z odpowiednio wyspecjalizowanych instytucji wydawniczych.

d) Badania nad sposobami oceniania uczniów przez nauczycieli trzeba by rozszerzyć na inne, stosowane narzędzia i techniki, aby się przekonać, czy zachodzi (może zachodzić?) w tym zakresie jakaś ich specjalizacja, tzn. które narzędzia i techniki najlepiej służą (mogą służyć?) pomiarowi jakiego rodzaju osiągnięć szkolnych.

e) Jeśli przyjąć, że sprawdziany pisemne trafnie ukazują sprawdzane przez polonistów osiągnięcia szkolne, to nie są to osiągnięcia, na których współczesnej szkole powinno najbardziej zależeć; i to zarówno w wymiarze kategorii taksonomicznych, jak i w wymiarze jakości materiału.

W wymiarze pierwszym dzieje się tak prawdopodobnie dlatego, że na ogół łatwiej sprawdza się to, co w taksometrii jest wartościowane najniżej.

W wymiarze drugim zaobserwowane zjawisko może mieć przyczynę w przestarzałej, nie modernizowanej wiedzy polonistycznej nauczycieli.

W obu przypadkach mamy też chyba do czynienia z niedostatkiem znajomości teorii nauczania, z zastępowaniem świadomości dydaktycznej — historycznoliteracką czy językoznawczą.

FROM RESEARCHES ON TEACHER'S WRITTEN TESTS ON POLISH STUDIES

* TADEUSZ PATRZAŁEK

Summary

In his paper the author describes his researches on teachers' tests on Polish studies, researches, he undertook in Poland for the firts time. The basic materials of the researches were tests published in our country in the years 1971 to 1975 with the purpose to give answers to three fundamental questions:

How often do teachers of Polish apply tests for estimation? What are their characteristics in comparison with the requirements of up-to-date test theories? Which kind of accomplishments of the pupils are tested?

Researches have shown, that in chosen Secondary Schools preparing for University, on an average avery fifth mark in the class register was given for tests and that teachers comparatively high estimate the dependability of these marks.

As for their arrangement tests are modifications of recapitulation tests, they are nearly exclusively composed of questions which need short answers, and they have numerous shortcomings and mistakes which often wreck their didactic value as a whole.

In the main they are directed on the lowest taxonomic categories, as memory or understanding and on material of relatively low value, according to modern didactics; among other things on formal periphrastics, as language and on context and world of the literary creation.

The researches bring to light the postulate for supplying the teacher with prepared and standardized tests elaborated by specialists and published by an appointed institute of nation-wide range.