

dr Henryk Palkij

Okręgowa Komisja Egzaminacyjna w Krakowie

Nowa formuła egzaminu maturalnego z historii i wiedzy o społeczeństwie – pierwsze doświadczenia

W 2015 roku zakończył się cykl wprowadzania zmian w systemie edukacji w Polsce spowodowany wprowadzeniem nowej *Podstawy programowej kształcenia ogólnego*¹. Objęły one nie tylko proces nauczania w szkołach, ale także system egzaminów zewnętrznych, gdyż wdrażana reforma wymusiła zmiany formuły egzaminów przewidzianych na wszystkich etapach nauczania. Wprowadzone zasady i arkusze miały duże znaczenie dla historii i wiedzy o społeczeństwie nauczanych w gimnazjum, bowiem zostały one w pewien sposób dowartościowane przez wprowadzenie odrębnej części egzaminu, poświęconej tym przedmiotom. Niewątpliwie podniosło to rangę obu przedmiotów w gimnazjum, a **w przyszłości pozwoli na poszerzenie zakresu prowadzonych badań nad efektami nauczania**². Pokazują to już pierwsze wyniki analiz umieszczane na stronach właściwych instytucji czy w różnych publikacjach³.

W przypadku egzaminu maturalnego z historii i wiedzy o społeczeństwie mamy inną sytuację. Egzamin w nowej formule przeprowadzony został po raz pierwszy, a uzyskane wyniki mogą stanowić postawę do wyciągania tylko ograniczonych wniosków i nie uprawniają także do szerszych porównań z wynikami z lat ubiegłych. Dopiero kolejne sesje egzaminacyjne pozwolą na pełniejsze określenie nowych tendencji, zarówno w odniesieniu do wyników egzaminu w poprzedniej formule, jak też w relacji do efektów nauczania w szkole według nowej *Podstawy programowej*.

Zmiany wprowadzone w obu egzaminach nie były zbyt radykalne, choć istotnie zmieniły charakter egzaminów z obu przedmiotów, co wydaje się być niedocenione⁴, mimo że było to już widoczne w propozycji nowych arkuszy⁵.

¹ Rozporządzenie MEN z 23 grudnia 2008, Dz. U. z dn. 15 stycznia 2009 nr 4, poz.17; *Podstawa programowa z komentarzami*, t. 4, *Edukacja historyczna i obywatelska*, MEN 2009, s. 34-48.

² Np. H. Szalenić, M. Grudniewska, B. Kondratek, F. Kulon, A. Pokropek, *Wyniki egzaminu gimnazjalnego z lat 2002-2010 na wspólnej skali*, „Edukacja” 2012, nr 3 (119) s. 9-30.

³ O wynikach osiągniętych w nowym kształcie egzaminu z historii i wiedzy o społeczeństwie pisał np. K. Jurek, *Egzamin gimnazjalny z historii. Obserwacje i wnioski* [w:] B. Niemierko i M.K. Szmigel (red.), *Regionalne i lokalne diagnozy edukacyjne*, Kraków 2012 s. 409-420. Ponadto można odesłać do licznych publikacji pojawiających się na stronach Centralnej Komisji Egzaminacyjnej oraz komisji okręgowych, jak również do publikacji z konferencji Polskiego Towarzystwa Diagnostyki Edukacyjnej z lat 2013 i 2014.

⁴ Informator umieszczony na stronie Centralnej Komisji Egzaminacyjnej: <http://www.cke.edu.pl/files/file/Matura-2015/Informatory-2015/Historia.pdf> oraz Wiedza o społeczeństwie [dostęp: 1.05.2014]; H. Palkij, *Egzamin maturalny z historii od 2015 roku*, „Hejnał Oświatowy” 2014, nr 3/131, s. 20-21; M. Kubala, *Egzamin maturalny z wiedzy o społeczeństwie od 2014/2015*, tamże s. 21-22.

⁵ Np. Arkusz maturalny z historii: http://www.cke.edu.pl/images/stories/0000000000000000002012_matura2012/hist/historia_pr.pdf [dostęp: 15.08.2013].

Oczywiście od dłuższego czasu toczyła się dyskusja nad kształtem egzaminów, tym bardziej że forma egzaminów istniejących od 2005 roku stale się zmieniała. W najnowszej zmianie wynikającej z nowej *Podstawy programowej* wykorzystano tylko niektóre propozycje, jakie pojawiały się w tej dyskusji⁶. Od momentu wdrożenia reformy w szkołach ponadgimnazjalnych utrzymywała się stale niepewność co do kształtu egzaminów, co wynikało z przyjętego sposobu pracy. W efekcie środowiska zainteresowane tym problemem miały ograniczony wpływ na ostateczny kształt wybranych rozwiązań, podobnie zresztą jak to było w przypadku przygotowania podstaw programowych, więc w praktyce na bieżąco musiały rozwiązywać pojawiające się problemy. Jest to jednak szerszy problem, który znacząco wychodzi poza ramy tego artykułu⁷.

Ze względu na duży zakres danych oraz krótki czas, jaki upłynął od egzaminów, oraz ograniczoną możliwość przeprowadzenia bardziej skomplikowanych analiz należy podkreślić wstępny charakter wniosków. Skorzystam z podstawowych danych z egzaminów przeprowadzonych przez Okręgową Komisję Egzaminacyjną w Krakowie⁸. Do analiz, które wykonał Przemysław Majkut, wykorzystano wyniki z podstawowych wersji arkuszy egzaminacyjnych.

Nowa struktura egzaminów z historii i wiedzy o społeczeństwie wymusza także ograniczenie porównań ze względu na liczne różnice. Na przykład w obu egzaminach zlikwidowano część testową sprawdzającą **opanowanie wiedzy**, **choć** **taka** **tendencja** **widoczna** **była** **od** **dawna** **w** **egzaminie** **maturalnym** **z** **historii**. **Ponadto** **zlikwidowano** **część** **źródłową**, **co** **szczególnie** **zmieniło** **arkusz** **z** **wiedzy** **o** **społeczeństwie**. Teraz arkusze egzaminacyjne z obu przedmiotów składają się z zadań opartych na analizie źródeł i wypracowań. Przy czym zwiększono ich liczbę – z historii mamy pięć wypracowań, w tym dwa z wyposażeniem źródłowym, natomiast z wiedzy o społeczeństwie zdający ma trzy wypracowania. Jednocześnie zmniejszono wagę punktową wypracowania w całym egzaminie – zamiast 20 punktów za wypracowanie zdający teraz otrzymuje 12 punktów. Ponadto za egzamin z historii zdający może otrzymać maksymalnie 50 punktów, a z wiedzy o społeczeństwie 60 punktów. W efekcie arkusze z obu egzaminów są zbyt obszerne, nawet na tle innych egzaminów maturalnych – zwłaszcza arkusz z historii, który liczył 37 stron.

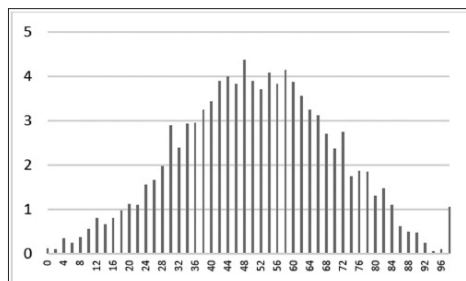
Co ważniejsze – oba przedmioty różni podejście do sposobu konstruowania poleceń, podejścia do sprawdzania wiedzy i umiejętności oraz sposób konstruowania kryteriów oceniania. Znacznie zmieniono kryteria oceniania wypracowań, zwłaszcza w egzaminie z wiedzy o społeczeństwie. Wszystkie te zmiany utrudniają porównanie wyników egzaminów z poprzednich sesji, a nawet porównanie wyników z obu przedmiotów.

⁶ Np. B. Andrzejewska, B. Freier-Pniok, *Ewaluacja arkuszy maturalnych z wiedzy o społeczeństwie i historii w zewnętrznym systemie oceniania* [w:] B. Niemierko i M.K. Szmigel (red.), *Ewaluacja w edukacji. Koncepcje, metody, perspektywy*, XVII Krajowa Konferencja Diagnostyki Edukacyjnej, Kraków, 23-25 września 2011, PTDE Kraków, s. 402-410.

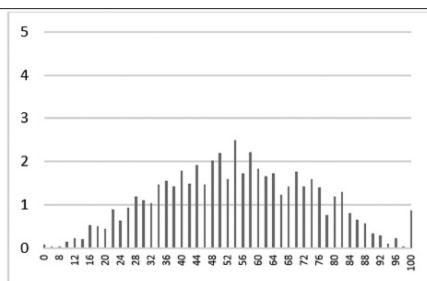
⁷ Pojawia się coraz liczniejsza literatura na ten temat, np. publikacje PTDE czy innych środowisk, jak: *Kierunki badań dydaktycznych. Kierunki zmian edukacji historycznej. X Toruńskie spotkania dydaktyczne* S. Roszak, M. Strzelecka, A. Wieczorek, Ł. Wróbel (red.), SOP Toruń 2014; liczne raporty różnych instytucji. IBE np. J. Chońska-Mika (red.), *Dobre praktyki w nauczaniu historii. Raport z badania jakościowego*, Warszawa, czerwiec 2014.

⁸ Wstępne wyniki egzaminów oraz skale centylowe opublikowano na stronie OKE w Krakowie: <http://www.oke.krakow.pl/inf/staticpages/index.php?page=20051012153316788> [dostęp: 1.07.2015].

Przechodząc do krótkiej analizy wyników egzaminu maturalnego z historii, należy wspomnieć, że ogólne wyniki były zbliżone do wyników osiągniętych w latach poprzednich. Najlepiej ilustrują to załączone rozkłady wyników z 2014 i 2015 roku.



Rysunek 1. Rozkład wyników z historii w 2014 roku



Rysunek 2. Rozkład wyników z historii w 2015 roku

Cechą charakterystyczną egzaminu maturalnego z historii była duża liczba zadań łatwych oraz bardzo łatwych, którą ilustruje złączona poniżej tabela. Na 31 zadań aż 12 osiągnęło wskaźnik łatwości wykonania powyżej 60%. Wydaje się, że wynika to z konstrukcji poleceń. Odnoszą się one najczęściej do dwóch lub trzech źródeł i sprawdzają często najbardziej podstawowe umiejętności, jak np. zadanie nr 1.1 i nr 1.2. Do poprawnego wykonania tych poleceń nie potrzeba nawet głębszej wiedzy historycznej. Czasem polecenia wymagają od zdającego tylko uzasadnienia tezy postawionej w poleceniu, a do uzasadnienia jej wszelkie informacje znajdują się w załączonych źródłach, np. zadanie 14. Oznacza to w istocie sprawdzanie podstawowych umiejętności czytania ze zrozumieniem oraz najprostszej selekcji informacji z tekstu bez wymagania znajomości szerszego kontekstu historycznego. Do tego typu zadań należą także zadania nr 10 oraz 21.1, w których zdający ma uzasadnić przedstawioną tezę.

Tabela 1. Wykonanie zadań z historii

Nr zadania	Procent wykonania	Średnia	Maksimum	Nr zadania	Procent wykonania	Średnia	Maksimum
R_z1.1	81,4	,814	1	R_z11	54,6	1,091	2
R_z1.2	87,8	1,755	2	R_z12	63,9	,639	1
R_z2	95,2	,952	1	R_z13.1	40,1	,401	1
R_z3.1	70,2	,702	1	R_z13.2	66,5	,665	1
R_z3.2	57,2	,572	1	R_z13.3	31,6	,316	1
R_z4	44,6	1,339	3	R_z14	79,9	,799	1
R_z5	40,2	,804	2	R_z15.1	14,4	,144	1
R_z6.1	27,6	,276	1	R_z15.2	41,0	,410	1
R_z6.2	65,5	,655	1	R_z16	54,7	,547	1
R_z7	53,2	,532	1	R_z17	58,7	1,175	2
R_z8.1	47,8	,478	1	R_z18	63,4	,634	1
R_z8.2	84,3	,843	1	R_z19	57,7	1,154	2

Nr zadania	Procent wykonania	Średnia	Maksimum	Nr zadania	Procent wykonania	Średnia	Maksimum
R_z8.3	53,6	,536	1	R_z20	66,5	,665	1
R_z9	55,7	,557	1	R_z21.1	83,1	,831	1
R_z10	80,5	1,610	2	R_z21.2	65,3	,653	1
				R_z22	24,2	2,906	12

Podaję te przykłady specjalnie, aby podkreślić jeden z ważniejszych problemów nowej formuły egzaminu – przeprowadzenie odpowiedniej selekcji umiejętności, które podlegają sprawdzeniu, oraz określeniu ich kontekstu historycznego. Ten ostatni problem wydaje się coraz bardziej palący, a jego pomyślnie rozwiązanie zdecydowanie o tym, na ile sprawdzane umiejętności będą dokładniej powiązane z konkretną wiedzą historyczną, a nie tylko z poziomem ogólnej wiedzy i inteligencji zdającego. Bez obudowania takich poleceń wiedzą historyczną zadania te nie będą różnicować populacji.

Z zadaniami tymi, jak też z zadaniami, w których podejmowana jest analiza źródeł ikonograficznych, np. zadania nr 7, 11, 13, 15, wiążą się jeszcze inne problemy. Pierwszym jest sposób formułowania poleceń. Rozbudowane polecenia, w których zdający wyszukuje właściwe elementy podlegające rozwinięciu w odpowiedzi, powodują, że w wielu przypadkach zdający nie potrafią odpowiedzieć prawidłowo. Ich wypowiedzi idą obok załączonych materiałów bądź są nieudolną próbą odtworzenia wiedzy związanej z załączonymi źródłami lub poleceniem, a nie właściwą analizą źródeł.


Na przykładzie stosunkowo prostego zadania nr 19 chciałbym wskazać kilka takich problemów⁹. Zdający otrzymał w wyposażeniu zadania plakat propagandowy, dane uzyskane w badaniach opinii publicznej i wiersz Mariana Żałuckiego oraz polecenie:

Wyjaśnij, dlaczego w 1980 r. można było uznać wiersz Mariana Żałuckiego za satyryczny komentarz do rezultatów polityki gospodarczej państwa z okresu tzw. Dekady Gierka.

Zdający powinien przeczytać wiersz i sformułować wyjaśnienie do tezy zawartej w poleceniu (tezę), odnieść ją do wiersza i do załączonych innych źródeł z dekady gierkowskiej. Model oceny zakładał wystawienie dwóch punktów za wyjaśnienie z odwołaniem się do dwóch źródeł, jednego punktu za odwołanie się do jednego źródła, natomiast zero punktów za odpowiedź błędną lub brak odpowiedzi.

⁹ Egzamin maturalny z historii. Poziom rozszerzony: <http://cke.edu.pl/index.php/2015-em> [dostęp: 1.07. 2015]

Źródło 1



M. Kallas, *historia ustroju Polski X-XX w.*, Warszawa 2001, il. VII.

Źródło 2. Komunikat Ośrodka Badania Opinii Publicznej i Studiów Programowych [fragment]

Ogólna ocena zaopatrzenia sklepów spożywczych znajdujących się w pobliżu miejsca zamieszkania badanych, dane w procentach

Zaopatrzenie	IV kwartał 1979 r.	III kwartał 1980 r
dobrze	10	4
takie sobie	29	18
złe	59	77
brak opinii	2	1

A. Dudek, Z. Zblewski, *Utopia nad Wisłą. Historia PRL-u*, Bielsko-Biala 2008, s. 286.

Źródło 3. Wiersz Mariana Załuckiego¹, *Przepis po polsku* opublikowany w Strajkowym Biuletynie Informacyjnym „Solidarność”, 25 sierpnia 1980 r.

Wziąć to, czego nie ma
dodać soli i kminku
potem zmieszać z tym czego
chwilowo brak na rynku

Mieszać długo i dokładnie
jak się znudzi przestać
i posypać tym na co
absolutnie nas nie stać [...]

Wszyscy u nas to jedzą
dla każdego wystarczy
na tym właśnie polega
Polski cud gospodarczy

Na podstawie: A. Dudek, Z. Zblewski, *Utopia nad Wisłą. Historia PRL-u*, Bielsko-Biala 2008, s. 277.

¹ Marian Załucki (1920-1979) – polski poeta, satyryk, pisarz.

Wyjaśnij, dlaczego w 1980 r. można było uznać wiersz Mariana Załuckiego za satyryczny komentarz do rezultatów polityki gospodarczej państwa z okresu tzw. dekady Gierka.

Przykładowa odpowiedź:

Wiersz Mariana Załuckiego opublikowany w Strajkowym Biuletynie Informacyjnym „Solidarność” w 1980 r. można było uznać za satyryczny komentarz do zakończonej fiaskiem polityki gospodarczej lat siedemdziesiątych. Według zapowiedzi władz jej celem miała być poprawa bytu mieszkańców – dostatnie życie [źródło 1.]. Błędna polityka doprowadziła do niedoborów w zaopatrzeniu w podstawowe artykuły spożywcze [źródło 2.], a ten problem poruszono w wierszu [zawierającym przepis na potrawę „z tego, czego nie ma” i mówiącego o rzekomym „cudzie gospodarczym”].

Zdający nie otrzymywali w poleceniu dokładniejszych wskazówek, jak zrealizować swoją odpowiedź, chociaż powinni to znać z praktyki szkolnej i zasad określonych w odpowiednich informatorach, więc rozwiązywali zadanie na różne sposoby.

Nr 1. Propaganda władzy czasów Gierka, której przykładem może być plakat ze źródła nr 1, twierdziła, że władze państwa mają na celu poprawienie sytuacji gospodarczej Polski i jej mieszkańców, co jednak ostatecznie okazało się fikcją, gdyż pod koniec tzw. dekady Gierka zaopatrzenie sklepów oceniano jako złe, a sytuacja ta pogarszała się, co zostało przedstawione w źródle 2. Wiersz Mariana Załuckiego mówi o wszechobecnych brakach w zaopatrzeniu i wykpiwa propagandę obozu rządzącego, który twierdził, że dzięki niemu doszło w kraju do „cudu gospodarczego”.

Nr 2. Ponieważ w polityce panowało przekonanie, że wszyscy robią wszystko dla dobra Polski, aby żyło się lepiej. Organizowano zjazdy itp. Jednak rzeczywistość była inna. W sklepach ciągle nic nie było, tylko podstawowe produkty. Do tego również nawiązuje Marian Załucki w swoim wierszu. Bo wszystko było dla każdego, dla każdego wystarczyło, lecz było to bardzo ubogie życie.

Nr 3. *Dekada Gierka była okresem rozwoju gospodarczego, ale prowadzonego na kredyt. Skutkowało to w następnej dekadzie ogromnym kryzysem gospodarczym i doprowadziło do importu żywności przez Polskę po raz pierwszy w historii. Wiersz Mariana Załuckiego można uznać jako satyryczny komentarz, ponieważ prześmiewczo pokazuje on w swoim utworze, że cud gospodarczy w dekadzie Gierka doprowadził Polskę do ruiny.*

Nr 4. *W 1980 roku można było uznać wiersz Mariana Załuckiego za satyryczny komentarz do sytuacji gospodarczej kraju, gdyż wyśmiewał on braki w sklepach i stosunkowo małe zarobki w porównaniu do cen niektórych produktów.*

Analizując liczne wypowiedzi, można było zauważyć, że zdający mieli różne problemy z prawidłową odpowiedzią. Często odpowiadali prosto na pytanie *dlaczego* na podstawie swojej wiedzy, nie odnosząc się w ogóle do źródeł (nawet to załączonego wiersza) lub odpowiadali w sposób ogólnikowy, nieporadny językowo. Inni odpowiadali, komentując wiersz M. Załuckiego, posługując się swoją wiedzą bez uwzględnienia innych źródeł (np. odpowiedź nr 3). Najczęściej odnoszono się do jednego lub drugiego źródła w sposób nieporadny, pozostawiając egzaminatorowi decyzję, czy może uznać sformułowanie zadającego za odniesienie się do źródła. Na przykład w odpowiedzi nr 4, w której nie wiadomo czy *braki w sklepach* odnoszą się tylko do wiersza, czy też do tabeli opisującej ocenę zaopatrzenia sklepów. U innych można znaleźć odniesienie się do źródeł (ich charakteru: propagandowego i oceniającego), jak np. w wypowiedzi nr 2. O wiele rzadsze były przypadki, w których mamy pełną wypowiedź.

Na podstawie tego zadania można pokazać kilka problemów. Pierwszym jest przygotowanie zdającego. Wyraźnie widać, że zdający nie zapoznali się z zasadami opisanymi w informatorach. Rozmowy ze zdającymi i nauczycielami ujawniły jeszcze inny problem, dosyć paradoksalny w świetle założeń reformy – brak czasu na pracę ze źródłami. Stąd brak umiejętności ich analizy, co zresztą potwierdzają inne zadania. Jest to sygnał rozchodzenia się kształtu egzaminu i praktyki szkolnej. Zdają się na to wskazywać także różne badania¹⁰, z których jednak wyciągane są odmienne wnioski dla praktyki egzaminacyjnej. Drugim problemem jest nieodpowiednie przygotowanie do egzaminu – brak znajomości typów poleceń, nieumiejętność odczytania polecenia, pogłębiająca się nieporadność językowa i brak wiedzy.

Inne problemy wywołuje konstrukcja arkusza, zadań oraz modelu odpowiedzi. Tutaj chciałbym podkreślić tylko kilka. Przed wszystkim widoczna jest nadmierna objętość egzaminu i ilości źródeł oraz częste sprawdzanie podobnych umiejętności na tym samym poziomie czy skomplikowane sformułowanie samych poleceń, co utrudnia w praktyce wskazanie, którą umiejętność sprawdzamy. Pojawia się wrażenie, że do prostej odpowiedzi tworzy się skomplikowaną obudowę, która służy tylko sformułowaniu prostego wniosku, bez posiadania większej wiedzy. Brak wiedzy powoduje pojawienie się najdziwniejszych wypowiedzi, często ahistorycznych, które dotąd występowały tylko w arkuszach podstawowych w początkowym okresie wprowadzenia

¹⁰ Np. raport IBE: J. Choińska-Mika (red.), *Dobre praktyki w nauczaniu historii. Raport z badania jakościowego*, Warszawa, czerwiec 2014.

egzaminów zewnętrznych, gdy zdający licznie wybierali historię. Jest to bardzo niepokojące teraz, w sytuacji gdy egzamin maturalny z historii wybierają osoby, które w szkole przeszły rozszerzony kurs z tego przedmiotu.

Taka sytuacja zwiększa trudność oceniania. Egzaminator otrzymuje „szeroki” klucz oceniania, w którym umieszczone są elementy ogólnie określone w poleceniu, i nie ma czytelnego progu zaliczenia. Z kolei zdający pisze bardzo nieporadnie, odnosząc się ogólnikowo do jednego z elementów polecenia, pogłębiając trudności w ocenie wypowiedzi. Przed egzaminatorem pojawia się pokusa uznania prawie każdej w miarę poprawnej wypowiedzi, chociaż celem polecenia było sprawdzenie innej umiejętności, a nie tylko poprawności merytorycznej. Zresztą także w tym przypadku często nie ma ostrych progów wyznaczających tę poprawność. W efekcie pojawia się tendencja uznawania wszelkich odpowiedzi (poza tymi oczywiście błędnymi). Jest to szczególnie widoczne w zadaniach, w których jest przyznawany jeden punkt, a takich jest większość, gdyż egzaminator ma bardzo mały margines na decyzję. W sumie prowadzi to do zawyżenia punktacji z całego arkusza.

Na problem ten można popatrzeć jeszcze w inny sposób, poprzez dokładne określenie kryteriów zaliczenia, ich „wyostrzenie”. W tym zadaniu mamy trzy źródła, z których wiersz stanowi podstawę do wyjaśnienia i argumentowania – zdający potwierdza tezę z polecenia (odwołuje się do wiersza i zgadza się z tezą) oraz w wyjaśnieniu czytelnie odwołuje się do źródeł (odnotowuje: źródło nr 1 – plakat czy też źródło nr 2 – tabela). Tak postawione kryteria spowodowałyby, że w cytowanych czterech odpowiedziach na zadanie 19 dwa punkty otrzymałby tylko zdający nr 1.

Jeśli w kryteriach określimy, że zdający „odnosi się do źródła”, a „odniesienie” pozostawiamy decyzji egzaminatora, dwa punkty otrzymuje zdający nr 2, gdyż egzaminator znajduje odniesienie do treści plakatu i tabeli – bez ich nazywania. Decyzja egzaminatora zależeć tutaj będzie z jednej strony od jego umiejętności zakwalifikowania wypowiedzi zdającego (czy on już „odnosi się”), czyli jego poziomu merytorycznego i mniej lub bardziej liberalnego podejścia do wypowiedzi zdającego. Z drugiej strony od umiejętności językowych, komunikacyjnych zdającego – od tego, w jaki sposób zdający sformułuje wypowiedź. Te umiejętności w szybkim tempie maleją. W ten sposób rozszerzamy krąg osób, które otrzymują maksymalną liczbę punktów. W obu przypadkach odpowiedzi 3 i 4 nie powinny być punktowane.

Jeśli spojrzymy literalnie na tak sformułowane polecenie (i uznamy, że zdający nie zna zasady zawartej w informatorze), to zauważymy, że do prawidłowej odpowiedzi wystarczyłoby nam odwołanie się do wiersza (źródło nr 3) i wiedza zdającego. Oznacza to, że można zakwalifikować jako poprawne także odpowiedzi nr 3 i 4. Egzaminator uznaje wtedy jednak, że wzmianka zdającego o wierszu Załuckiego oznacza odniesienie się do źródła, natomiast wyjaśnienie polega na skorzystaniu z wiedzy własnej, którą egzaminator bardzo „luźno” uznaje za odniesienie do pozostałych źródeł. Czasem zdarzały się takie błędy egzaminatorów.

Moim zdaniem taka sytuacja ujawnia głębszy problem. W opisanych pierwszych dwóch przypadkach, chociaż o różnej skali ostrości czy kategoryczności, sprawdzamy umiejętność pracy ze źródłem. Wydaje się więc nieuzasadnionym

brak wyraźnego polecenia do uwzględnienia w odpowiedzi zdającego źródeł nr 1 i 2. Rodzi to niepewność nie tylko w skali poprawności stosowania kryteriów (literalnie czy liberalnie), ale skłania egzaminatorów do uznawania za poprawne wspomnianych odpowiedzi nr 3 i 4.

Nie oznacza to, że musimy rezygnować z rozwiązania drugiego, czyli uzasadnienia własną wiedzą tezy postawionej w poleceniu i wierszu Załuckiego. Oznacza to jedna, że sprawdzamy inne umiejętności, np. orientację zdającego o epoce Gierka. Wówczas do poprawnego rozwiązania zadania wystarczy nam tylko wiersz (źródło nr 3) i polecenie, więc oszczędzamy wiele miejsca, pozbywając się pozostałych dwóch źródeł.

Kiedy podczas procesu oceniania istnieją jednocześnie obie postawy, zaburza to wyniki oraz analizę tych wyników, bowiem nie wiemy właściwie, jakie umiejętności zostały sprawdzone i ocenione. W efekcie egzaminator podświadomie posługuje się dwoma rodzajami kryteriów, a zdający rozwiązują jakby dwa różne zadania, które sprawdzają różne umiejętności. Na koniec u egzaminatora i zdającego pojawia się wrażenie, że w tym zadaniu można otrzymać punkty za każdą w miarę poprawną merytorycznie (zgodnie z wiedzą historyczną) odpowiedź, chociaż w istocie to nie jest prawda, bo ocenione są jakby dwa zadania. Nie znamy skali tego zjawiska, gdyż jest ono niedocenione i nie prowadzono szczegółowych analiz. Najwyraźniej ujawnia się ono wówczas, gdy informacje zawarte w źródłach są sprzeczne z wiedzą historyczną. Często zdający nie potrafią wtedy skonfrontować źródła i piszą „obok” lub wbrew informacjom i poleceniom zawartym w zadaniu.

Wydaje się, że obserwacje powyższe potwierdza częściowo analiza wypracowań (zadanie 22). W nowej formule pojawiło się pięć tematów, w tym dwa z wyposażeniem źródłowym, które zdający powinni uwzględnić w swej wypowiedzi.

1. Załamanie się imperium rzymskiego w V w. mogłoby nastąpić nawet bez interwencji barbarzyńskich najeźdźców – Christopher Dawson. Scharakteryzuj wpływ czynników wewnętrznych i zewnętrznych na upadek cesarstwa zachodniorzymskiego i ustosunkuj się do zacytowanej tezy historyka brytyjskiego.
2. Wyjaśnij wpływ uwarunkowań zewnętrznych i lokalnych na gospodarkę polską w okresie rozbitcia dzielnicowego. W pracy odwołaj się do materiałów źródłowych (str. 26–28).
3. Rozstrzygnij, czy reformacja stanowiła czynnik dezintegrujący społeczeństwo monarchii polsko-litewskiej za panowania ostatnich Jagiellonów i w Rzeczypospolitej Obojga Narodów w drugiej połowie XVI i pierwszej połowie XVII w. Odpowiedź uzasadnij.
4. Czy zgadzasz się z opinią opublikowaną w XIX w. na łamach emigracyjnego pisma „Baczość”, że *rewolucja społeczna, ażeby była narodową – potrzebuje powstania. Powstanie, żeby żyć – potrzebuje rewolucji społecznej?* Odpowiedź uzasadnij. W pracy odwołaj się do materiałów źródłowych (str. 28–29).
5. Anglia miała wybór pomiędzy wojną i hańbą. Wybrała hańbę, ale będzie miała *wojnę* – Winston Churchill. Scharakteryzuj i oceń stanowisko Wielkiej Brytanii wobec polityki III Rzeszy w latach 1933–1940 i ustosunkuj się do wypowiedzi Winstona Churchilla z 1938 r.

W wyborach zdających dominowały najnowsze epoki (XIX i XX wiek). Temat z epoki nowożytnej został w praktyce pominięty, i faktycznie był trudny. Można zauważyć, że wprowadzenie tematu ze źródłami (tematy nr 2 i 4) nie spowodowało jakiegos znaczącego skoku wyborów przez zdających. Kryterium wyboru była raczej tematyka niż forma zadania.

Tabela 2. Rozkład wyboru tematów przez zdających

Temat 1	Temat 2	Temat 3	Temat 4	Temat 5
19,3%	15,8%	3,5%	31,6%	29,6%

Wypracowania oceniane były według zmodyfikowanych kryteriów z zachowaniem holistycznego oceniania, przy zmniejszeniu liczby punktów (do 12), z zachowaniem poziomów, ale ze zmniejszoną i zróżnicowaną skalą punktową na poszczególnych poziomach. 0 punktów otrzymywały prace świadczące o zupełnym braku zrozumienia tematu lub takie, w których nie podjęto próby rozwiązania. Utrzymano cztery poziomy (I skala 1-2 pkt, II skala 3-5 pkt, III skala 6-8 pkt, IV 9-12 pkt).

Interesujące okazały się konsekwencje wprowadzenia takiego rozwiązania, które prezentują poniższe tabele zawierające rozkłady wyników na poszczególne tematy. Pomimo ułatwienia, jakie dawała większa liczba tematów, średni wynik za wypracowanie okazał się zaskakująco niski czyli 3 pkt na 12 możliwych. Wyniki zdających wybierających różne tematy była zbliżona za wypracowanie, przy czym tematy z wyposażeniem źródłowym okazały się trudniejsze. Średni wynik zdających według tematów, które wybrali, przedstawiono w tabeli.

Tabela 3. Wyniki wykonania zadania 22 według tematów wybranych przez zdających

	Temat 1	Temat 2	Temat 3	Temat 4	Temat 5
Średnia w pkt	3,50	2,44	3,45	2,45	3,54
% punktów	28	20	20	20	30

Jeśli przyjrzymy się procentowemu rozkładowi wyników w tabeli 4., to wyraźnie widać nowe zjawiska. Zaliczam do ich: rosnący udział prac ocenionych na 0 punktów, bardzo duży procent prac o niskich wynikach (1-2 pkt) oraz niewiele prac z najwyższą punktacją (11-12 pkt). Przy tym szczególnie źle wypadają tematy z wyposażeniem źródłowym.

Tabela 4. Procentowy rozkład częstości wystąpienia punktacji w zadaniu 22.

Liczba punktów	Temat 1	Temat 2	Temat 3	Temat 4	Temat 5
	procent	procent	procent	procent	procent
0	7,6	9,5	14,7	10,0	8,4
1	19,9	35,3	21,6	32,2	22,8
2	20,8	22,9	18,6	22,3	18,0
3	11,2	7,8	3,9	11,2	8,8
4	8,3	6,1	8,8	10,1	10,7
5	9,1	7,1	3,9	5,0	7,3
6	7,8	4,5	10,8	3,6	6,6
7	5,2	4,5	6,9	1,6	4,8

Liczba punktów	Temat 1	Temat 2	Temat 3	Temat 4	Temat 5
	procent	procent	procent	procent	procent
8	5,3	,6	3,9	1,6	4,9
9	2,0	,2	2,0	1,0	2,8
10	2,1	1,1	3,9	,9	2,8
11	,4	0	1,0	,4	1,9
12	,4	,2	0	,1	,3
Ogółem	100,0	100,0	100,0	100,0	100,0

Widać to wyraźnie w przypadku rozdziału wyników na poszczególne poziomy (tabele pokazujące rozkład liczebności wyników oraz wykresy z rozkładem procentowym wyników). Przy czym należy pamiętać, że do poziomów nie wliczono prac ocenionych na 0 punktów (są poza nimi), a poziom pierwszy ma skalę dwupunktową, gdy natomiast poziom najwyższy, czwarty – skalę czteropunktową. Trzeba o tym pamiętać zwłaszcza przy porównywaniu poziomów z ubiegłych lat.

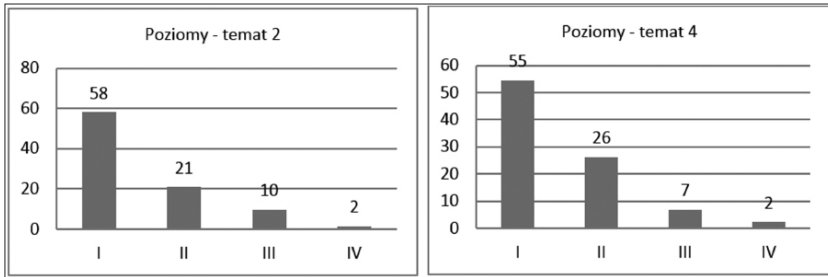
Tabela 5. Rozkład liczebności wyników w podziale na poszczególne tematy

Poziom	Temat 1	Temat 2	Temat 3	Temat 4	Temat 5
0	43	44	15	92	72
I	229	269	41	502	351
II	161	97	17	242	231
III	103	45	22	63	140
IV	27	7	7	22	67
Ogółem	563	462	102	921	861

W sumie wyniki uzyskane w każdym z wypracowań pokazują znaczące przesunięcie w dół w porównaniu z poprzednimi latami. Jeśli do poziomu I dodamy prace ocenione na 0 pkt, to okaże się, że zdający, którzy otrzymali 0-2 pkt, stanowią od 50% do ponad 60% piszących egzamin. Jeśli z drugiej strony zauważymy, że najwyższy, IV poziom ma największą skalę punktów, a w poszczególnych tematach osiągnęło go tylko od 2% do 8% zdających, to można uznać, że mamy problem, który można nazwać znaczącym spadkiem umiejętności sformułowania samodzielnej dłuższej wypowiedzi, przy zaniżonych wymaganiach merytorycznych. Oczywiście może to budzić niepokój. Zauważalne jest wielkie rozwarstwienie wyników, są prace bardzo słabe, przeciętne i niewielka liczba prac dobrych i bardzo dobrych. Trzeba się zastanowić, jakie czynniki spowodowały takie skutki, o ile kolejne sesje egzaminacyjne potwierdzą te prawidłowości.

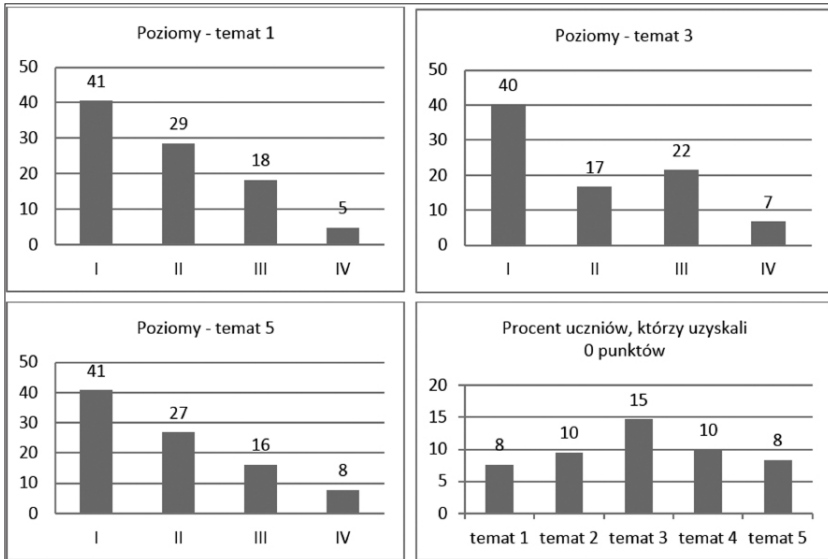
Analizując procentowe rozkłady wyników na poszczególne poziomy (bez prac ocenionych na 0 pkt), możemy zaobserwować kilka zjawisk. Wyraźnie widać porażkę osób wybierających zadanie z wyposażeniem źródłowym; większość z nich, bo prawie 80%, a gdy dodamy prace na 0 pkt, to prawie 90 procent zmieściła się w I i II poziomie (do 5 pkt). Zjawisko to potwierdza bardziej szczegółowa analiza prac. Zdający ograniczali się do opisu źródeł, który często był na najniższym poziomie, czyli opisu tego, co jest w źródłach. Nie posiadali umiejętności połączenia źródeł z wiedzą własną. Pomijam już sposób formułowania myśli. Na tym etapie trudno jednak pokusić się o dokładniejsze wnioski.

Niewątpliwie jednak należy zastanowić się nad celowością stosowania takiej formy wypracowania i kryteriami jego oceny.



Rysunek 3. Procentowy rozkład wyników według poziomów dla tematów z załączonymi źródłami

Wypracowania „zwykle”, polegające na stworzeniu własnej wypowiedzi na zadany problem, mają lepsze wyniki, a właściwie „strukturę” wyników, czyli procentowy rozkład na poziomy. Wyraźnie odstaje od nich wypracowanie trzecie – ma największy odsetek niskich wyników (licząc zadania z uzyskanym wynikiem 0 pkt) oraz zaburzony rozkład wyników w poziomie III. Świadczy to nie tylko o trudności problemu, przed jakim stanęli zdający, ale chyba także o problemach egzaminatorów oceniających to zadanie.

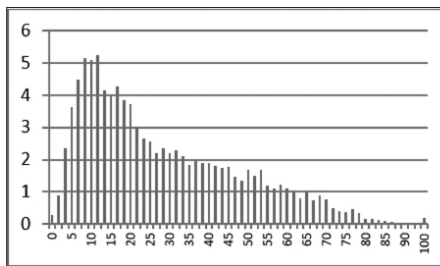


Rysunek 4. Procentowy rozkład wyników według poziomów dla tematów tradycyjnych, czyli dla wypracowań

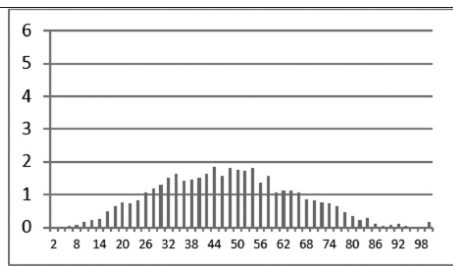
W sumie pierwsza sesja egzaminacyjna dostarcza ciekawego pola do obserwacji efektów wprowadzonych zmian w systemie nauczania historii i sposobie sprawdzania jej wyników. Jest jednak za wcześnie na formułowanie zdecydowanych wniosków.

Przed nową sytuacją stajemy także w przypadku **egzaminu maturalnego z wiedzy o społeczeństwie**. Przed jego analizą należy jednak wspomnieć, że nowa *Podstawa programowa* znacząco zmieniła pozycję tego przedmiotu. Teraz, dla osób, które wybrały WOS jako egzamin maturalny, jest on o wiele trudniejszy i nauczany w szerszym zakresie merytorycznym i czasowym niż poprzednio. Wydaje się, że wielu zdających nie zauważyło tej zmiany, a raczej zlekceważyło jej konsekwencje. Do tej sytuacji dołożyła się zmianą formuły egzaminu maturalnego i kryteriów oceniania – zwłaszcza wypracowań.

W efekcie wynik całego arkusza okazał się zaskakująco niski, około 28%, wobec zwykle uzyskiwanych około 46%. Do tego należy zauważyć znaczącą zmianę w rozkładzie wyników w porównaniu z innymi latami. Pokazuję ją **poniżej**, pomimo że porównanie takie obarczone jest licznymi pułapkami. Jednak w tym przypadku bardzo dobrze pokazują one wielką polaryzację wyników. Z jednej strony istnieje dominacja wyników niskich, a z drugiej mamy niewielką liczbę wyników wysokich. Na to zjawisko niewątpliwie ma wpływ poziom nauczania w szkołach, poziom kadry nauczycielskiej i chyba poziom trudności przedmiotu.



Rysunek 5 Rozkład wyników z WOS-u w 2015 roku



Rysunek 6. Rozkład wyników z WOS-u w 2014 roku

Analizując wybrane wskaźniki statystyczne poszczególnych zadań znajdujących się w arkuszu, **trzeba zauważyć, że są tylko dwa zadania łatwe powyżej 60% wykonania**, a 10 zadań poniżej 20% wykonania i 12 zadań poniżej 30% wykonania, czyli razem stanowią one ponad połowę ogólnej liczby zadań. Przy czym wstępna analiza merytoryczna poszczególnych zadań nie wskazuje na to, aby faktycznie w arkuszu znajdowały się zadania szczególnie trudne w odniesieniu do tematyki poruszanej na lekcjach wiedzy o społeczeństwie.

Można stąd wysunąć wniosek, że na taki wynik złożyło się słabe przygotowanie merytoryczne zdających, osobnym problemem jest to, z czego ono wynika, forma arkusza – sposób formułowania poleceń oraz stosowany schemat oceniania. Już wstępna analiza zadań pokazuje, że sformułowane są one trochę inaczej niż zadania znajdujące się w arkuszu historii. Polecenia sprawdzają w większym stopniu wiedzę zdającego oraz wybrane, pojedyncze umiejętności. Ponadto polecenia, które odnoszą się do załączonych materiałów źródłowych, nastawione są na bardziej jednoznaczne sprawdzenie wybranych umiejętności, najczęściej powiązanych z wiedzą merytoryczną, a w niewielkim stopniu pozwalają na poprawną odpowiedź wynikającą tylko z prostego wyciągnięcia informacji z załączonego źródła. Analiza źródeł jest dokładniej ukierunkowana, raczej odnosi

się do bardziej szczegółowych problemów, nie polega tylko na uzyskaniu ogólniejszych sformułowań i wniosków zdającego. Do poleceń „źródłowych” często dodawane są zadania łączące wiedzę o problemie ze ściślejszym powiązaniem z materiałem źródłowym. Mało jest zadań, w **których zdający formułuje ogólniejsze opinie lub ocenia ogólniejsze zjawiska**. Przez to zadania te wymagają od zdającego konkretnej wiedzy, która sięga nie tylko do ogólnych zjawisk, ale nawet w niektóre niższe poziomy oraz skupienia się **podczas formułowania własnych opinii**. Ponadto zdanie utrudnia zdającym fakt skomplikowanej struktury WOS-u. Składa się on w praktyce z różnych przedmiotów, więc zdającym często brakuje odpowiedniego aparatu pojęciowego, który umożliwiłby im poprawne sformułowanie własnej opinii. Postępujący spadek umiejętności komunikacji językowej pogarsza jeszcze ich sytuację.

Sposób konstrukcji poleceń powoduje także konsekwencje w sposobie oceniania. Model i kryteria zaliczenia odpowiedzi stają się bardziej jednoznaczne, więc egzaminator pewniej przydziela punkty, gdyż poziom niepewności ogranicza się właściwie do sfery językowej zdającego – oczywiście poza problemami, które z natury rzeczy są niejednoznaczne lub w różny sposób opisywane w literaturze przedmiotu. Nie można zapominać także, że **arkusz ma 60 punktów do uzyskania**, czyli o 10 punktów więcej niż arkusz egzaminacyjny z historii. Zwiększając liczbę zadań w naturalny sposób można sprawdzić większy zakres treści kształcenia, co na pewno ma wpływ na zwiększenie skali trudności egzaminu.

Tabela 6. Wykonanie zadań z wiedzy o społeczeństwie

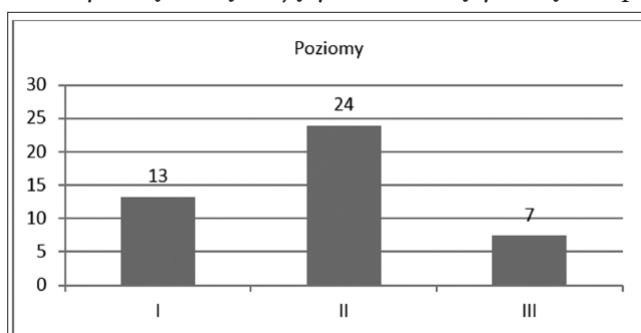
Numer zadania	Procent wykonania	Średnia	Punkty	Numer zadania	Procent wykonania	Średnia	Punkty
R_z1	38,2	,382	1	R_z21	19,5	,195	1
R_z2	25,5	,511	2	R_z22	26,2	,262	1
R_z3	50,7	,507	1	R_z23	30,6	,306	1
R_z4	39,9	,399	1	R_z24	37,2	,744	2
R_z5	41,1	,411	1	R_z25	27,3	,546	2
R_z6	29,3	,293	1	R_z26	32,8	,328	1
R_z7	49,7	,497	1	R_z27	17,6	,352	2
R_z8	36,2	,362	1	R_z28	65,2	,652	1
R_z9	28,0	,280	1	R_z29	51,9	,519	1
R_z10	15,7	,157	1	R_z30	9,7	,097	1
R_z11	30,7	,307	1	R_z31	29,5	,884	3
R_z12	35,6	,356	1	R_z32	63,2	,632	1
R_z13	11,7	,234	2	R_z33	21,5	,215	1
R_z14	27,0	,270	1	R_z34	18,5	,370	2
R_z15	22,1	,221	1	R_z35	11,1	,111	1
R_z16	48,1	,481	1	R_z36	43,9	,439	1
R_z17	18,8	,188	1	R_z37	12,3	,246	2
R_z18	24,2	,242	1	R_z38	48,4	,484	1
R_z19	32,2	,322	1	R_z39	18,6	,186	1
R_z20	22,8	,228	1	R_z40	22,4	2,693	12

Widać to szczególnie w przypadku wypracowań. Także w przypadku WOS-u nastąpiło zwiększenie ich liczby i zmniejszenie liczby punktów. Niestety z powodu konstrukcji karty odpowiedzi nie można podjąć próby analizy poszczególnych wypracowań. Musimy zadowolić się ogólnymi wnioskami odnoszącymi się wspólnie do wszystkich wypracowań, co znacznie zubaża wnioskowanie. Także w przypadku WOS-u zdający, który uzyskali 0 punktów, znaleźli się poza skalą poziomów. Jeśli popatrzymy na częstości i procent skumulowany, to okazuje się, że prace na 0 punktów stanowiły aż 55,4% wszystkich prac. Taki wynik świadczy o wielkim rozwarstwieniu na tych, którzy cokolwiek wiedzą albo chcą napisać, i tych, którzy podjęli próbę rozwiązania przedstawionego problemu. Źródła tej sytuacji muszą podlegać w przyszłości bardzo dokładnej analizie.

Tabela 7. Rozkład częstości wyników z wypracowań z wiedzy o społeczeństwie

	Częstość	Procent	Procent ważnych	Procent skumulowany
0	3502	55,4	55,4	55,4
1	61	1,0	1,0	56,3
2	72	1,1	1,1	57,5
3	278	4,4	4,4	61,8
4	428	6,8	6,8	68,6
5	377	6,0	6,0	74,6
6	447	7,1	7,1	81,6
7	564	8,9	8,9	90,5
8	129	2,0	2,0	92,6
9	123	1,9	1,9	94,5
10	190	3,0	3,0	97,5
11	120	1,9	1,9	99,4
12	36	,6	,6	100,0
Ogółem	6327	100,0	100,0	

Jeśli popatrzymy natomiast na wyniki składające się na poszczególne poziomy, to zauważmy, że układają się one odmiennie niż w przypadku wypracowań z historii. Widzimy dużą liczbę zdających mieszczących się w II poziomie.



Rysunek 7. Procentowy rozkład wyników w podziale na poziomy z wiedzy o społeczeństwie

W przypadku wiedzy o społeczeństwie właściwie niemożliwe jest porównywanie z poprzednimi latami wyników wypracowań. Wynika to nie tylko ze zmniejszenia liczby punktów za wypracowanie – do 12, ale głównie z zupełnie odmiennej konstrukcji zasad oceniania. Wprowadzone kategorie: wartość merytoryczna, z przydzielonymi 4, 8, 12 punktami, stanowią poziom, z których wychodzi egzaminator. Powoduje to odmienne rozumienie pojęcia „poziom”, gdyż korygowane jest ono przez kryteria: poprawność merytoryczna, selekcja informacji oraz język, styl i kompozycja, które mają zasadniczy wpływ na ostateczny wynik, jaki uzyskuje piszący wypracowanie. Ponadto w kategorii wartość merytoryczna dodatkowo bardziej sztywno i szczegółowo wydzielone są poszczególne kryteria merytoryczne, które zdający powinien spełnić.

W efekcie mamy dziwną sytuację. Jeśli zdający otrzymał od egzaminatora 6 punktów, to może być tak, że jako wartość wyjściową otrzymał 8 punktów (wartość merytoryczną na poziomie II), ale niepoprawna selekcja oraz błędy językowe spowodowały zmniejszenie liczby punktów do 6. Może być też odwrotnie. Zdający może mieć 4 punkty (wartość merytoryczna na poziomie I), ale poprawność merytoryczna, selekcja lub język zostały ocenione na poziomie II bądź wyższym, więc spowodowało to, że otrzymał ostatecznie 6 lub 7 punktów. W ten sposób mamy paradoks – dodatkowe kryteria: poprawność merytoryczna, selekcja i język mogą powodować, że praca, w której temat został zrealizowany przez zdającego słabo pod względem merytorycznym, może być wyżej oceniona niż praca, w której realizacja tematu jest bardzo dobra, a której punktacja może być obniżona przez te dodatkowe kryteria.

W efekcie dodatkowe kryteria: selekcja, poprawność merytoryczna i język mogą „fałszować wyniki”, gdyż informacje dodatkowe świadczące o dużej wiedzy lub dodatkowe, ale błędne mogą zostać potraktowane jako elementy obciążające, powodujące, że przez te trzy kryteria egzaminator może zaniżyć wynik w pracach bardzo dobrych, a podnieść go w pracach przeciętnych. Może to także wynikać z efektu egzaminatora. Uważam, że problem należy zbadać na większej liczbie prac i przy pomocy specjalnej metodologii – z użyciem próby kontrolnej. Na efekt egzaminatora składa się prawdopodobnie nadmierna ostrość w przypadku błędów językowych i selekcji, zwłaszcza przy dobrych pracach, oraz problem zrozumienia podstawowych zapisów klucza, zwłaszcza w kryterium selekcja i język. Natomiast kryteria: wartość merytoryczna (dzięki bardziej szczegółowym zapisom) oraz poprawność merytoryczna wydają się łatwiejsze do stosowania przez egzaminatora.

Uwzględniając te uwagi, spróbowałem pokazać, jaki istotnie wpływ na „poziom” mogą mieć te wszystkie kryteria. Kolorami zaznaczono „zakres poziomu”, który powstaje wówczas, gdy egzaminator wychodzi przy ocenianiu z poszczególnych wartości punktowych przypisanych kryterium wartość merytoryczna. Widzimy, że „poziomy” mają zdecydowanie inny zakres oraz przyjmują odmienną wartość punktową.

Tabela 8. Rozkład częstości wyników z wypracowań z wiedzy o społeczeństwie

Punkty	Częstość	Procent	Częstość	Procent	Częstość	Procent
0	3502	55,4	3502	55,4	3502	55,4
1	61	1,0				
2	72	1,1				
3	278	4,4				
4	428	6,8				
5	377	6,0	377	6,0		
6	447	7,1	447	7,1		
7	564	8,9	564	8,9		
8			129	2,0		
9			123	1,9	123	1,9
10			190	3,0	190	3,0
11			120	1,9	120	1,9
12					36	,6
Ogółem	6327	100,0	6327	100,0	6327	100,0

Podsumowując pierwszą sesję egzaminu maturalnego z historii i wiedzy o społeczeństwie przeprowadzoną według nowych zasad opartych na nowej *Podstawie programowej*, należy stwierdzić, że stoimy przed nowymi problemami, które trzeba będzie poddać szczegółowej analizie w przyszłych sesjach egzaminacyjnych, aby lepiej zrozumieć zjawiska zachodzące w obu przedmiotach. Teraz jesteśmy dopiero na początku drogi, a więc jest to dobry moment na refleksję i wnioski oraz sprawdzenie, czy wyniki analizy w skali kraju będą podobne.