

# Efekt zwrotny (washback) – próba opisu zjawiska<sup>1</sup>

Elżbieta Zawadowska-Kittel  
Lingwistyczna Szkoła Wyższa w Warszawie

## Abstract

This article ventures to explicate the notion of *washback/backwash* as well as its synonymous concept *efekt zwrotny* employed in academic and critical works on the subject by English-speaking and Polish researchers, provide the outline of the concept in question and finally attempt to analyse it. As the notions of *washback/backwash* have been assimilated in Polish as borrowings, they are likely to be utilised in academic publications together with the Polish counterpart – *efekt zwrotny*. Hence, the present author continues to refer to these terms when citing individual scholars' standpoints. Although the notion of washback and its role in teaching/learning process resists unanimous interpretations, its definitions provoke no heated debates. Broadly, scholars are in agreement that it denotes the influence of testing on teaching and learning process, however, when attempting to describe this influence, they are either neutral – aware of both assets and drawbacks of the phenomenon – or else they take sides by seeking to underscore its positive or ill effects. Multifaceted nature of washback inevitably triggers critical discussions on its range, meaning and character. This article offers a round-up of conceptualisations and types of washback, typology of the notion as well as description of its derivatives. Also, the present author's own definition of washback, covering its major distinguishing features has been proposed.

**Keywords:** washback (backwash), high-stakes examinations, learning/teaching process

---

<sup>1</sup> Artykuł stanowi część większego opracowania.

**Abstrakt**

Celem artykułu jest wyjaśnienie terminu efekt zwrotny oraz terminów do niego synonimicznych (*backwash/washback*) używanych w literaturze przedmiotu przez anglojęzycznych i polskich badaczy zjawiska oraz jego charakterystyka i próba analizy. Przeniesienie terminów *washback/backwash* do języka polskiego jest typowym zapożyczeniem i można przewidywać, że terminy *washback* i *backwash* będą nadal funkcjonować w literaturze przedmiotu obok terminu efekt zwrotny i dlatego nie rezygnujemy z użycia tych terminów przy okazji cytowania poglądów poszczególnych badaczy. Zjawisko efektu zwrotnego nie jest interpretowane jednoznacznie, choć jego definicje nie budzą kontrowersji – ogólnie rzecz biorąc, badacze są zgodni co do tego, iż jest to wpływ testowania na proces nauczania i uczenia się, przy czym, opisując ów wpływ, zachowują albo postawę neutralną, dostrzegając zarówno negatywny, jak i pozytywny charakter tego zjawiska, lub też eksponują jego pożądane lub niepożądane skutki. Wieloaspektowość zjawiska efektu zwrotnego wywołuje również nieuniknioną polemikę co do jego zasięgu, znaczenia i charakteru. Artykuł zawiera przegląd koncepcji oraz propozycje modeli efektu zwrotnego, a także typologię zjawiska oraz charakterystykę jego pochodnych. Przedstawiono w nim również własną definicję efektu zwrotnego, uwzględniającą najważniejsze jego cechy charakterystyczne.

**Słowa kluczowe:** *washback* (*backwash*), efekt zwrotny, egzamin doniosły, proces nauczania/uczenia się

Celem artykułu jest wyjaśnienie terminu *efekt zwrotny* oraz terminów do niego synonimicznych (*backwash/washback*) używanych w literaturze przedmiotu przez anglojęzycznych i polskich badaczy zjawiska oraz jego charakterystyka i próba analizy.

Zjawisko efektu zwrotnego nie jest interpretowane jednoznacznie, choć jego definicje nie budzą kontrowersji – ogólnie rzecz biorąc, badacze są zgodni co do tego, iż jest to wpływ testowania na proces nauczania i uczenia się, przy czym, opisując ów wpływ, zachowują albo postawę neutralną, dostrzegając zarówno negatywny, jak i pozytywny charakter tego zjawiska, lub też eksponują jego pożądane lub niepożądane skutki. Wieloaspektowość zjawiska efektu zwrotnego wywołuje również nieuniknioną polemikę co do jego zasięgu, znaczenia i charakteru.

Artykuł zawiera przegląd koncepcji oraz propozycje modeli efektu zwrotnego, a także typologię zjawiska oraz charakterystykę jego pochodnych. Przedstawiono w nim również własną definicję efektu zwrotnego, uwzględniającą najważniejsze jego cechy charakterystyczne.

W literaturze przedmiotu dominują odniesienia do publikacji anglojęzycznych, ponieważ w Polsce nie prowadzono jak dotychczas szeroko zakrojonych badań tego zjawiska. Co ciekawe, niektórzy badacze polscy nie używają nawet w swoich publikacjach terminu *efekt zwrotny* i posługują się terminem angielskim *washback* lub *backwash*, ewentualnie *washback effect/backwash effect*, wybranym z literatury przedmiotu raczej przypadkowo. Warto w tym miejscu zaznaczyć, że wszystkie te terminy powstały równolegle do siebie i były używane alternatywnie, choć terminy *washback/washback effect/backwash effect* używane są wyłącznie przez badaczy zjawiska. W słownikach występuje jednak wyłącznie termin *backwash* – „the bad situation that remains after something that has happened” (Longman Dictionary of Contemporary English), czyli „niekorzystna sytuacja, jaka następuje w wyniku zaistniałego wydarzenia” lub „the unwelcome repercussions of some social action” (New Webster Comprehensive Dictionary), czyli „nieprzyjemne skutki zdarzenia lub sytuacji”. W słownikach bilingwalnych pojęcie to wyjaśniane jest jako: „wycofywanie się fali; niemiłe następstwa” (Słownik angielsko-polski Exeter), „wir wodny za płynącym statkiem” (Wilson). Tak czy inaczej, definicje terminu używanego w języku potocznym podkreślają pejoratywny charakter zjawiska, innymi słowy – oplakane skutki czegoś, co się zdarzyło.

Takie negatywne konotacje nie znajdują jednak logicznego odzwierciedlenia w terminologii używanej przez badaczy zjawiska omawianego w artykule, ponieważ w kontekście dydaktycznym termin *backwash* może mieć zarówno pozytywny, jak negatywny charakter. W artykule generalnie przyjmujemy polski termin *efekt zwrotny*, podając jednak również terminologię angielską, tam gdzie ma to naszym zdaniem znaczenie dla historii użycia tego terminu, zwłaszcza że termin *washback* jest ściśle związany ze zjawiskiem wzajemnego wpływu, jaki wywierają na siebie testy oraz proces uczenia się i nauczania, i przed rozpoczęciem badań na temat tego zjawiska w ogóle w języku nie istniał.

Przeniesienie terminów *washback/backwash* do języka polskiego jest typowym zapożyczeniem i można przewidywać, że terminy *washback* i *backwash* będą nadal funkcjonować w literaturze przedmiotu obok terminu *efekt zwrotny* i dlatego nie rezygnujemy z użycia tych terminów przy okazji cytowania poglądów poszczególnych badaczy.

## Ustalenia terminologiczne

Zjawisko wpływu sposobu testowania istniało zawsze, stanowi jednak przedmiot dociekań i badań dopiero od niedawna. Pierwsze próby określenia i zdefiniowania tego wpływu nastąpiły w latach dziewięćdziesiątych, wtórnie do koncepcji „nauczania sterowanego

pomiarem” Williama Pophama, która zakładała, że sposób testowania i egzaminowania powinien stanowić siłę napędową nauczania, a zatem również uczenia się.

Opisując zjawisko polegające na tym, że sposób testowania wpływa na proces nauczania, J. Charles Alderson i Diane Wall (115) używają terminu *washback* i taki właśnie termin przeważa w dydaktyce brytyjskiej. John Biggs („Assumptions” 12) wybiera jednak termin *backwash*, stwierdzając, że testowanie wpływa nie tylko na program nauczania, ale również na metody nauczania i podejście uczących się do procesu uczenia się. J. Biggs („The psychology of assessment” 11) wyróżnia przy tym sześć funkcji testów kluczowych: 1) kontrola programów kształcenia, 2) likwidacja korupcji, 3) stworzenie motywacji do osiągnięcia większych kompetencji językowych, 4) narzędzie do selekcjonowania najzdolniejszych, 5) pomiar jakości procesu uczenia się i nauczania, 6) ograniczenie różnic w programach nauczania (w krajach, gdzie podstawy programowe, czyli przyjęty na danym etapie nauczania zestaw treści programowych, nie podlegają centralizacji). Definiując w ten sposób efekt zwrotny, zakłada raczej jego pozytywną rolę w edukacji.

Bernard Spolsky natomiast sugeruje (55), że termin *backwash* powinien być używany w odniesieniu do przypadkowych czyli niezamierzonych efektów ubocznych egzaminów, a nie konsekwencji sytuacji, w której zakładanym, świadomym celem nadrzędnym egzaminów jest uzyskanie kontroli nad programem nauczania. B. Spolsky przyznaje również, że nie znajduje żadnego uzasadnienia leksykalnego dla terminu *washback* ukutego przez dydaktyków.

Liying Cheng (“How does washback influence teaching?” 36) połączyła te koncepcje, stwierdzając, że *washback* to zamierzony kierunek i funkcja zmian programowych osiągane poprzez modyfikacje w systemie egzaminów państwowych.

Mimo różnic w użytej terminologii żaden z badaczy nie negował wpływu testowania na proces nauczania/uczenia się. Nieco inaczej określano jednak sfery tych wpływów; w zależności od interpretacji wpływowi ulegają obszary węższe lub szersze. Mary Finocchiaro i Sydney Sako (311) podkreślają wpływ sposobu testowania na wybór treści kształcenia i procedur nauczania. Shaun Gates pisze po prostu o wpływie testowania na „nauczanie i uczenie się”. Lyle Bachman (283) odnosi się jednak tylko do wpływu egzaminów na nauczanie.

Harold S. Madsen (120) jest zdania, że czasem skupienie się nauczyciela na słownictwie lub gramatyce może zaowocować pozytywnym wpływem (*backwash*) na proces nauczania. Zawęża zatem definicję wyłącznie do pozytywnego oddziaływania tego zjawiska na nauczanie.

Istnieją również poglądy skupiające się na wybranych, pojedynczych aspektach problemu efektu zwrotnego. Przykładem może tu być Philip Shawcross, który wyraża pragmatyczną opinię, twierdząc, że definicja efektu zwrotnego (*washback*) zawiera się w krótkim pytaniu: „Czy to będzie na egzaminie?”.

Prekursorem myśli, że prawidłowo skonstruowane egzaminy wywierają pozytywny wpływ na nauczanie, a egzaminy nieprzemyślane mogą zniszczyć ten proces, jest Alan Davies („The Construction...” 42), przekonany, że testy wywierają wpływ zarówno na program nauczania, jak i samo nauczanie i dlatego twórcy testów powinni zadbać o to, by był to wpływ pozytywny.

Sama świadomość wpływu, jaki testy wywierają na proces nauczania/uczenia się nie jest zatem nowa. Cytowana wcześniej D. Wall, współpracująca z J. Ch. Aldersonem, w swojej samodzielnej publikacji (291) dokonuje jednak rozróżnienia pomiędzy wpływem (*impact*) i efektem zwrotnym (*washback*, *backwash*). Według Diane Wall wpływ odnosi się do wszelkiego rodzaju efektów, jaki test może wywierać na jednostki, politykę edukacyjną, system edukacyjny, procedury stosowane w szkołach bądź klasach czy nawet na społeczeństwo jako całość, natomiast *washback* (*backwash*) to wpływ testu wyłącznie na proces nauczania/uczenia się.

Gary Buck (17) zakłada, że istnieje naturalna tendencja zarówno wśród nauczycieli, jak i uczących się, aby dopasowywać aktywności klasowe do wymagań testu, zwłaszcza gdy od jego wyniku zależy przyszłość zdających. Taki właśnie wpływ testu na to, co się dzieje w klasie (*washback*), jest niezwykle istotny i może być zarówno dobroczynny, jak i bardzo szkodliwy.

Arthur Hughes („Testing” 1) stwierdza, że „wpływ testu na proces uczenia się i nauczania” należy określać jako *backwash* i wskazuje warunki, jakie muszą zostać spełnione, by ów wpływ był pozytywny. Wylicza wśród nich m.in. położenie akcentu na testowanie umiejętności, które w pierwszym rzędzie chcemy rozwijać, ustalenie ścisłych kryteriów testowania, jawność testów (nauczyciele i uczniowie wiedzą, co i w jaki sposób będzie testowane), opieranie testów osiągnięć na jasno sformułowanych celach nauczania, tzw. testowanie bezpośrednie, przeprowadzenie testów pilotażowych na dużej reprezentatywnej grupie uczących się, pomoc dla nauczycieli przygotowujących do testu. Koncepcja Artura Hughesa omawiana jest szerzej w dalszej części artykułu.

Trzeba jednak zauważyć, że wspomniana przez A. Hughesa jawność może wpłynąć na nauczanie zdecydowanie negatywnie, ponieważ prowadzi do uczenia „pod test”, co może skutkować nawet ograniczeniem programu nauczania, a już na pewno wybiórczym traktowaniem treści programowych oraz selektywnym podejściem do aktywności

lekcyjnych i ćwiczeń językowych. Bardzo często bowiem typ zadań wykonywanych na lekcji stanowi odzwierciedlenie zadań egzaminacyjnych, co z pewnością ogranicza możliwości rozwoju ucznia, a także zabija w dużym stopniu inwencję nauczyciela, który czuje się w pierwszym rzędzie zobligowany, aby przygotować uczniów do egzaminu stanowiącego o ich przyszłości. Tego typu egzaminy nazywamy za Bolesławem Niemierko egzaminami doniosłymi.<sup>2</sup>

Opisując „wykorzystanie zewnętrznych testów językowych do wywierania wpływu i kierowania procesem uczenia się i nauczania”, Elana Shohamy (513) również używa terminu *washback*. Podobnie jak Gary Buck podkreśla, że zjawisko to jest ściśle związane z dużym znaczeniem egzaminów doniosłych i ich wpływem na przyszłość zdających. E. Shohamy (wraz z in.) dodaje, że moc sprawcza i znaczenie egzaminów umożliwiają wykorzystanie ich jako narzędzi kontrolujących system edukacji i kształtujących zachowania tych, których przyszłość zależy w jakimś stopniu od wyników testów.

Według E. Shohamy („The Power of Tests”) efekty egzaminów doniosłych obejmują cały kompleks zjawisk związanych ze zmianami zachowań, jakie zachodzą w wyniku egzaminu. Tak pojęty efekt egzaminów na proces nauczania może mieć charakter społeczny lub edukacyjny. E. Shohamy posługuje się szerokim pojęciem *efekt*, na które składa się efekt zwrotny (*washback*) i konsekwencje (*consequences*) testu. Autorka ta nie utożsamia konsekwencji z efektem zwrotnym, gdyż stanowią one obok niego część całego zjawiska, jakim jest oddziaływanie testu na proces uczenia się i nauczania. Przykładami efektów edukacyjnych mogą być programy nauczania, metody nauczania, strategie uczenia się, nowe materiały dydaktyczne, sposób oceniania. Efekt społeczny to oddziaływanie testu na takie aspekty życia, jak ideologia, zasady etyczne i moralność.

E. Shohamy uważa, że użytkowe podejście do testów wykracza poza granice dydaktyki. Przyjąwszy ten punkt widzenia, należy oczekiwać ogromnej odpowiedzialności od decydentów, czyli osób, od których zależy polityka edukacyjna oraz od autorów tych zmian. A to oznacza, że praca nad testem trwa również po jego wprowadzeniu, wyniki są dogłębnie analizowane, a badania statystyczne podawane do wiadomości wszystkich zainteresowanych. Ponadto należałoby badać, jak dany egzamin doniosły wpływa na poszczególne jednostki.<sup>3</sup>

Równie ważny aspekt zagadnienia to wpływ, jaki na test wywierają czynniki pozadydaktyczne, tj. polityczne, kulturowe czy społeczne. Test jest bowiem „zanurzony”

<sup>2</sup> Artykuł poświęcony egzaminom doniosłym ukazał się w I numerze Rocznika LSW.

<sup>3</sup> Tego typu próby podejmuje pisząca te słowa w badaniach nad efektem zwrotnym Nowej Matury na proces uczenia się i nauczania.

w konkretnym kontekście edukacyjnym, politycznym i kulturowym, w którym różne siły dochodzą do głosu, niejednokrotnie ze sobą walcząc.

W przeciwieństwie do tradycyjnego, dydaktycznego rozumienia testu, podejście użytkowe zakłada zbadanie różnych aspektów problemu, informowanie społeczeństwa o wynikach i ewentualnych konsekwencjach niewłaściwego wykorzystania wyników testu.

Tego typu podejście do testowania ma zdecydowanie bardziej humanistyczny charakter, dzięki czemu lepiej służy systemowi i społeczeństwu, w którym przeprowadza się testy.

Samuel Messick („Validity and Washback” 241) stwierdza wręcz, że efekt zwrotny testu (*washback*) to stopień, w jakim test wpływa na zachowanie uczących się, zmuszając ich do wykonania działań, których nie podjęliby się zapewne w sytuacji, gdyby nie istniał test.

Do testu postrzeganego jako katalizator zmian nawiązuje również Ian Pearson, który uważa, że istnieje tendencja, aby egzaminy państwowe traktować jako dźwignię zmian wpływającą na modyfikacje postaw zarówno nauczycieli, jak i uczniów, oraz katalizującą powstawanie nowych aktywności lekcyjnych.

Należy zauważyć, że najbardziej istotne dla zrozumienia roli testowania w procesie edukacyjnym jest ujęcie analizujące wpływ egzaminów zarówno dla procesu nauczania, jak i uczenia się, gdyż stanowią one integralną całość przedstawioną – jak piszą Weronika Wilczyńska i Anna Michońska-Stadnik (31) – w skrótowym zapisie N/U J2, uwydatniającym „fundamentalną dla naszej dziedziny współzależność między tym, co można określić jako dwie strony procesu glottodydaktycznego”. Dlatego najpełniejsza wydaje się interpretacja A. Hughesa („Backwash”), w której analizowany jest wpływ testu na oba procesy, choć jak zauważono wcześniej, niektóre jego wnioski są co najmniej kontrowersyjne. Zauważyć też należy, że szerokie koncepcje zjawiska uwzględniające wpływ testowania zarówno na proces nauczania, jak i uczenia się wymagają dokładnego omówienia, dokonanego w dalszej części artykułu.

Podejścia węższe, obejmujące jedynie poszczególne aspekty zagadnienia, są jednak również istotne, ponieważ każde z nich stanowi ważny, oryginalny przyczynek do całościowej analizy zagadnienia.

Naszym zdaniem zbadanie wpływu każdego nowego egzaminu, zwłaszcza doniesłego na proces uczenia się i nauczania może znacząco wpłynąć na poszerzenie wiedzy o zjawisku efektu zwrotnego o nowe elementy, jakie mogą wynikać ze specyfiki samego testowania, grupy poddanej egzaminowi i znaczenia tego testu dla dalszych losów zdających. Następuje zatem sprzężenie pomiędzy wiedzą teoretyczną na temat zjawiska efektu zwrotnego i badaniami empirycznymi. Z jednej strony bowiem założenia teoretyczne czy hipotezy wymagają potwierdzenia w praktyce i zapewne niektóre z nich mogą

zostać pozytywnie zweryfikowane w pewnych określonych warunkach i sfalsyfikowane w innych, z drugiej zaś – dzięki badaniom empirycznym wpływu nowych egzaminów na proces nauczania i uczenia się można uzupełnić wiedzę teoretyczną na ten temat o nowe aspekty zjawiska.

Dokonania badaczy anglojęzycznych na tym polu wydają się bardziej znaczące z uwagi na fakt, że tym większy jest wpływ testu na proces nauczania i uczenia się, im bardziej istotny jest wpływ testu na losy zdających. W przypadku scentralizowanych zewnętrznych egzaminów doniosłych może być to wpływ stanowiący o ich dalszej przyszłości edukacyjnej, a taki właśnie system testowania przyjęto w Wielkiej Brytanii, a następnie skopiowano na przykład w Chinach i Hongkongu (stąd badania Xi Li nad wpływem matury z języka angielskiego na proces uczenia się i nauczania oraz praca Li Cheng skoncentrowana wokół egzaminu maturalnego w Hongkongu). Ponadto badacze anglojęzyczni (np. Charles Alderson i Liz Hamp-Lyons) zajmowali się również wpływem prestiżowych międzynarodowych testów języków (TOEFL) na proces uczenia się i nauczania (Luxia 50–51).

Można zatem stwierdzić, że zainteresowanie badaczy anglojęzycznych wpływem testowania na proces uczenia się i nauczania wynika z tradycji edukacyjnych tych krajów i adaptowania podobnych systemów ewaluacyjnych w innych krajach w Europie i na świecie. Adaptacja takiego systemu może być w niektórych przypadkach uzasadniona historycznie (Hongkong), w innych jest sprawą wyboru władz oświatowych.

W Polsce znaczny wkład w dociekania związane z efektem zwrotnym ma Bolesław Niemierko, który stwierdza, że wpływ egzaminu doniosłego, czyli takiego, w którym znaczenie informacji o wyniku uczenia się jest większe od komentarza dydaktycznego (ang. *high-stake test*, czyli *egzamin wysokich stawek*) na treść i formę kształcenia, nazywamy jego efektem zwrotnym. B. Niemierko dodaje, że ten efekt może być niszczący, gdyż zarówno uczniowie, jak i rodzice mogą żądać, by większość czasu lekcyjnego poświęcić na przygotowanie do egzaminu, co skutkowałoby pomijaniem wszystkiego, co nie będzie testowane (265).

Hanna Komorowska stwierdza, że sposób prowadzenia kontroli i oceny zawsze wpływa na proces uczenia się i nauczania. Podzielając pogląd L. Cheng, przypomina, że wpływ zwrotny testu na metody nauczania, określane jako *washback effect*, jest znacznie bardziej skuteczny od zaleceń programowych czy wskazówek metodycznych.

O efekcie zwrotnym testu wspomina również Maria Wysocka w kontekście badanych przez nią różnych stylów dydaktycznych reprezentowanych przez polskich nauczycieli. Pisząc o wpływie testu (*backwash effect*), stwierdza, że: „niepokojąco często proces dydaktyczny prowadzony przez badanych nauczycieli był – jak to sygnalizowali – zdominowany przez czekające uczniów egzaminy”, i dodaje, iż najczęściej był to egzamin doj-



rzałości. M. Wysocka zwraca uwagę, że zorganizowanie całego procesu nauczania pod kątem sprawdzianu musi oddziaływać negatywnie na jego ostateczny kształt i stanowi czynnik hamujący pracę nauczyciela nad sobą, gdyż: „skoncentrowany jest on na egzaminie (którego zakres i forma jest podana do wiadomości), nie zaś na przekazaniu uczniom określonej wiedzy, z której nieznany wcześniej fragment będzie w nieznany im wcześniej sposób egzaminem objęty”.

Paweł Poszytek (46) za Cameronem J. Weirem posługuje się pojęciem *washback validity*, omawiając trafność testu – „washback validity concerns the issue of how tests influence teaching and learning”. Trafność efektu zwrotnego jest zatem związana z tym, w jaki sposób testy wpływają na proces uczenia się i nauczania.

Jak wspomniano wyżej, badacze zjawiska eksponują jego różne aspekty. Aby podsumować te dokonania, spróbujmy przyjąć definicję uwzględniającą najważniejsze dokonania w tym zakresie.

Efekt zwrotny egzaminów jest to zatem wpływ egzaminów doniosłych, zamierzony bądź nie, zarówno pozytywny, jak i negatywny na szeroko rozumiany proces uczenia się i nauczania oraz na wszystkich jego uczestników, a nawet na cały system edukacji.

Naszym zdaniem nie tyle jednak wyniki testów, lecz znaczenie, jakie przykłada do nich społeczeństwo, a zwłaszcza uczestnicy procesu edukacji, stanowi o sile efektu zwrotnego testu. Sam wynik testu, wyrażony procentowo czy też za pomocą oceny, która pozostaje bez konsekwencji dla dalszych losów zdającego, ma znaczenie jedynie informacyjne, lecz nie selektywne (wynik testu nie ogranicza w żadnym stopniu kariery edukacyjnej czy zawodowej zdających, nie wyklucza ich z grona potencjalnych pracowników czy studentów). Dlatego dopiero instrumentalizacja takich wyników stanowi groźne narzędzie w rękach decydentów, od których należy wymagać ogromnej odpowiedzialności, jaka może się wyrażać również w sposobie opisywania rezultatów (czy są wyrażone jedynie procentowo lub za pomocą oceny, jak ma to miejsce w przypadku egzaminów Cambridge). Skłaniamy się zdecydowanie ku komunikowaniu ocen w formie opisowej, ponieważ ich analiza może służyć egzaminowanemu jako wskazówka na przyszłość, a ewaluacja zmienia charakter z czysto sumującego na kształtujący.

## Próba analizy zjawiska efektu zwrotnego

Efekt zwrotny (*washback*) to zjawisko wielowymiarowe, a zatem – by w pełni zrozumieć jego rolę w procesie uczenia się i nauczania – należy je rozważać kompleksowo, z uwzględnieniem wszystkich aspektów i pojęć, jakie są z nim nierozzerwalnie związane.

Takie pojmowanie problemu wymaga odniesień do tych koncepcji zjawiska, które wpisują efekt zwrotny w sferę trafności testu, jak czynią to S. Messick („Validity”, „Validity and Washback”) czy John R. Frederiksen i Allan Collins, jak i do poglądów reprezentowanych m.in. przez Ch. Aldersona i D. Wall lub A. Daviesa („The Construction”), którzy rozpatrują ten problem niezależnie od jego trafności.

Aby zrozumieć złożoność omawianego zjawiska, należy się zapoznać zarówno z jego typologią, jak i mechanizmami jego działania. Przedstawione poniżej koncepcje stanowiły ważne osiągnięcia w badaniach nad efektem zwrotnym.

### Hipotezy Ch. Aldersona i D. Wall

Swoisty przełom w badaniach nad efektem zwrotnym stanowią hipotezy Ch. Aldersona i D. Wall (120–121) ukute na podstawie ich wniosku, że definicje używane poprzednio nie są wystarczająco jasne, gdyż wynikają raczej z pewnych założeń teoretycznych niż badań empirycznych. Ch. Alderson i D. Wall przedstawiają zatem cały szereg hipotez ilustrujących wpływ, jaki testy mogą wywrzeć na proces uczenia się i nauczania. Dowodzą, że autorzy testów powinni określić, jaki cel zamierzają osiągnąć poprzez ten a nie inny typ testu oraz uświadomić sobie dokładnie, jaki efekt zwrotny (*washback*) zamierzają wywrzeć poprzez dany test. Oto zaprezentowany przez nich zbiór 15 hipotez:

1. Test wpływa na proces nauczania.
2. Test wpływa na proces uczenia się.
3. Test kształtuje treści nauczania.
4. Test kształtuje sposób nauczania.
5. Test wpływa na to, czego uczą się uczniowie.
6. Test wpływa na to, jak uczą się uczniowie.
7. Test wpływa na jakość nauczania i kolejność wprowadzania kolejnych elementów.
8. Test wpływa na jakość uczenia się i kolejność przyswajania kolejnych elementów.
9. Test wpływa na poziom nauczania i jego głębię.
10. Test wpływa na poziom uczenia się i jego głębię.
11. Test wpływa na podejście do treści nauczania, na stosowane metody i techniki nauczania.
12. Testy istotne dla ucznia zawsze wywierają wpływ na proces uczenia się i nauczania.
13. Testy mało istotne nie wywierają wpływu na proces uczenia się i nauczania.
14. Testy wywierają wpływ na wszystkich uczących się i nauczających.
15. Testy wywierają wpływ tylko na niektórych uczących się.

Interesujące jest to, że niektóre z hipotez postawionych przez Ch. Aldersona i D. Wall mają przynajmniej potencjalnie charakter komplementarny (1 i 2, 3 i 4, 5 i 6, 7 i 8, 9 i 10), a niektóre wzajemnie się wykluczają. Hipotezy 5 i 6 (czego i jak uczą się uczniowie), 12 i 13 powinny mieć przynajmniej teoretycznie zdecydowanie komplementarny charakter, gdyż nowe programy nauczania najczęściej wymagają nowych technik nauczania. W rzeczywistości nowe techniki nauczania nie zawsze są jednak stosowane przy zmianie sposobów testowania (Zawadowska-Kittel). Hipoteza o wpływie egzaminu na to, czego uczą się uczniowie, dotyczy głównie treści (często wprowadzanych w nowych podręcznikach). Może się jednak zdarzyć, że nauczyciel realizuje podręcznik, nie stosując nowych metod i technik nauczania. Z badań Jo Lewkowicz i Elżbiety Zawadowskiej („Classroom-based assessment”) nad sposobami testowania w Polsce na lekcjach języków obcych w szkołach podstawowych, gimnazjach i liceach wynika, że nauczyciele często stosują technikę tłumaczenia zdań z J1 na J2 i z J2 na J1, choć podręczniki przez nich realizowane mają już charakter komunikacyjny, a co najważniejsze – komunikacyjny charakter mają testy gimnazjalne i maturalne. Dlatego również wyłącznie teoretycznie należałoby rozważać komplementarny charakter hipotezy 3 i 4. Można jednak założyć, że jakość nauczania wpływa na jakość uczenia się, zatem hipotezy 7 i 8 zdecydowanie się ze sobą łączą, podobnie jak hipotezy 9 i 10 są komplementarne z tych samych powodów. Dopełnianie się hipotez 12 i 13 jest oczywiste, gdyż obie uzależniają stopień, w jaki dany test wpływa na nauczających i uczących się, od znaczenia testu.

Jeśli zaś chodzi o hipotezy 14 i 15, to te pozornie wykluczające się naszym zdaniem hipotezy mogą się jednak mimo wszystko również uzupełniać: w pewnych przypadkach z badań empirycznych może wynikać prawdziwość hipotezy 14, w innych – 15. Wydaje się jednak, że testy doniosłe determinują sposób uczenia się i nauczania w sposób zasadniczy, toteż można założyć, że w przypadku większości testów egzaminu wpływają na wszystkich uczących się i nauczających. W wyniku takiego założenia otrzymalibyśmy hipotezę 15b, której udowodnienie byłoby jednak bardzo trudne logistycznie, gdyż wymagałoby badań o charakterze światowym na niespotykaną dotąd skalę.

Generalnie rzecz biorąc, wydaje się, że na hipotezach Ch. Aldersona i D. Wall można oprzeć każde badania empiryczne.

Do omawianych powyżej hipotez Ch. Aldersona i D. Wall dodano (Alderson, Hamp-Lyons) jeszcze hipotezę 16, która zakłada, że:

16. Testy wywierają inny rodzaj wpływu na różne grupy uczących i uczących się.

Oznacza to, że siła efektu zwrotnego i jego rodzaj są różne w zależności od statusu testu, różnic pomiędzy sposobem testowania a dominującym w danej chwili sposobem nauczania,

stopnia, w jakim autorom testów i nauczającym zależy na odpowiednim sposobie przygotowywania zdających do danego testu oraz ich woli wprowadzania nowego testu w życie.

Jest to założenie niezmiernie ciekawe, zwłaszcza jeśli weźmiemy pod uwagę kryteria, na podstawie których można by tworzyć profile takich grup. W przypadku uczących się jedynymi z ważniejszych kryteriów byłyby zapewne wiek badanych i charakter selekcji egzaminu, lecz równie ważnymi mogłyby się okazać ich miejsce zamieszkania, status społeczny rodziny, plany na przyszłość i inne. W przypadku nauczających mogłoby to być zapewne ich wykształcenie, doświadczenie, wiek i być może bardziej szczegółowe kryteria, takie jak np. znajomość skali biegłości językowej stosowanej w *Europejskim systemie opisu kształcenia językowego* (ESOKJ) (Lewkowicz, Zawadowska „Classroom-based assessment”).

Sprawdzenie trafności hipotez to droga trudna, żmudna i daleka, konieczna jednak do przebycia, gdyż tylko dzięki niej można wysnuć dalej idące wnioski na temat procesu kształcenia. Co więcej, na polskim przykładzie (Lewkowicz, Zawadowska-Kittel „Impact”; Zawadowska-Kittel) widać, że pewne poglądy na temat tych samych egzaminów ewoluują w czasie. Ponadto ewoluują egzaminy; można również próbować wykazać związek przyczynowo-skutkowy między tą ewolucją i zmianami opinii na ich temat.

## Koncepcja Arthura Hughesa

Arthur Hughes („Backwash”), który obok Ch. Aldersona i D. Wall uczestniczył w programie badań prowadzonych przed zmianą egzaminu *TOEFL*, rozróżnia typy efektu zwrotnego na oddziaływanie na uczestników procesu, procesy i produkty systemu edukacji. W tej trychotomii uczestnikami procesu są nauczyciele, uczniowie, władze edukacyjne, autorzy materiałów dydaktycznych i wydawcy. Wprowadzenie nowych testów wpływa na ich sposób myślenia oraz sposób działania, w przypadku autorów podręczników i wydawców wpływa również zarówno na ilość, jak i na jakość publikowanych materiałów dydaktycznych. A. Hughes podkreśla, że aby taki wpływ wystąpił, uczniom i nauczycielom powinno zależeć na pozytywnym wyniku testu, a uczestnicy procesu nauczania i uczenia się muszą posiadać niezbędną wiedzę na temat konstrukcji testu oraz kryteriów oceny. Muszą również mieć dostęp do materiałów umożliwiających przygotowanie do testu.

Wprowadzenie rozróżnienia A. Hughesa wydaje się niezwykle cenne, ponieważ umożliwia ono uporządkowanie analizy efektu zwrotnego, a także – już z racji swego istnienia – zakłada jego zróżnicowany charakter.

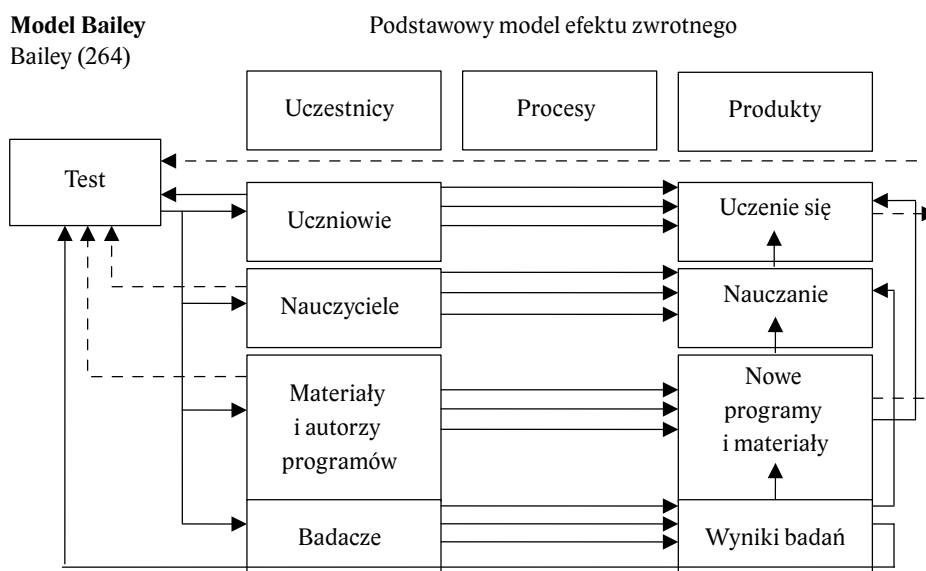
Druga część poglądów A. Hughesa dotycząca wiedzy uczniów na temat konstrukcji testu, kryteriów oceny oraz dostępu do materiałów umożliwiających przygotowanie do testu budzi wątpliwości. Oczywiście format egzaminu i kryteria jego oce-

ny powinny być podawane do wiadomości, lecz zbyt szczegółowe upowszechnianie takich danych łączy się z ryzykiem nauczania „pod test” i wykluczeniem z programów nauczania elementów nietestowanych, a w konsekwencji z ich zawężeniem. Jeśli chodzi o materiały przygotowujące do testu, to według Williama Mehrensa i Johna Kaminsky’ego nauczanie sprawności według formatu zadań przypominających zadania testowe stanowi rodzaj swoistej granicy pomiędzy zachowaniami etycznymi i nieetycznymi, natomiast nauczanie oparte na publikowanych testach tego samego formatu, a zwłaszcza upublicznienie go wśród uczniów jest zdecydowanie nieetyczne, wręcz naganne.

Naszym zdaniem zbyt precyzyjne podawanie formatu testu i rozpowszechnianie zadań egzaminacyjnych musi przyczynić się do wzrostu ryzyka nauczania „pod test”, które jest związane z całym szeregiem negatywnych skutków omawianych w dalszej części artykułu.

## Model Kathleen Bailey

Połączenie hipotez badawczych Ch. Aldersona i D. Wall z koncepcją A. Hughesa stanowi model Kathleen Bailey (264), która dokonała próby zdefiniowania efektu zwrotnego, sposobu jego oddziaływania, a także metod, jakie można wdrożyć do procesu nauczania/uczenia się, by wpływ ten był wyłącznie pozytywny. Konstruuąc własny model, wskazała również, w jaki sposób można to zjawisko badać. W swojej koncepcji K. Bailey rozróżnia uczestników, procesy i produkty efektu zwrotnego, wskazując na ich wzajemne zależności.



Rysunek 1. Model Kathleen Bailey

Odnosząc się do hipotez Ch. Aldersona i D. Wall, K. Bailey dzieli je na dwie podgrupy. Jak wskazuje D. Wall („The impact of high-stakes examinations”), pierwszą z grup (hipotezy 2, 5, 6, 8, 10, 11) odnosi do efektu zwrotnego, jaki test wywiera na uczących się, i podaje przykłady działań podejmowanych przez uczących się w ramach przygotowań do testu. Są to m.in.: wykonywanie zadań podobnych formalnie do zadań testowych, ćwiczenie strategii wykonywania testu, uczestnictwo w kursach przygotowujących do testu, aż po opuszczanie regularnych zajęć lekcyjnych po to, by w tym czasie przygotowywać się do testu (264–265). K. Bailey przypisała drugą część hipotez (1, 3, 4, 7, 9) do efektu zwrotnego, jaki test wywiera na program nauczania.

Należy jednak zauważyć, że równie zasadny byłby tu podział tych hipotez na te, które odnoszą się do nauczających (1, 3, 4, 7, 9, 11, 13, 14), i te odnoszące się do uczniów (2, 5, 6, 8, 10, 12, 13, 14, 15).

Zważywszy, że proces uczenia się i nauczania są ze sobą nierozzerwalnie związane, można stwierdzić, że hipotezy odnoszące się do sposobu nauczania (1, 3, 4, 7, 9, 11) odnoszą się również do uczniów. Hipotezy 12, 13, 14 już z założenia odwołują się zarówno do nauczających, jak i uczących się.

K. Bailey sformułowała również szereg pytań, jakie powinni postawić sobie twórcy nowo wprowadzanego testu, by zyskać pewność, że efekt zwrotny, jaki wywrze test na wszystkich jego uczestnikach, będzie pozytywny. Pytania te są następujące:

1. Czy uczestnicy procesu rozumieją cel testu i zamierzone skutki jego wprowadzenia?
2. Czy wyniki testu przekazywane są uczestnikom procesu w sposób jasny i zrozumiały?
3. Czy wyniki testu są postrzegane przez uczestników jako wiarygodne i sprawiedliwe?
4. Czy test sprawdza dokładnie treści zawarte w programie?
5. Czy cele testu są wystarczająco jasne?
6. Czy test opiera się na rozsądnych przesłankach teoretycznych znajdujących wiarygodne potwierdzenie w prowadzonych badaniach empirycznych?
7. Czy w teście wykorzystywane są teksty autentyczne i zadania zbliżone do autentycznych działań językowych?
8. Czy uczestnicy procesu są zaangażowani w proces oceniania? (276–277)

Pytania 1 i 5 odnoszą się do celu testu i jego zamierzonych skutków, 2, 3 i 8 – do jego oceniania i wyników, natomiast pytania 4, 6 i 7 związane są bezpośrednio ze sposobem konstruowania zadań. Należy zauważyć, że jeśli chodzi o efekt zwrotny, K. Bailey zajmuje się wyłącznie efektem zwrotnym zamierzonym, pomijając kwestię „skutków ubocznych”, a te właśnie są zwykle najbardziej zgubne. Należałoby zatem dodać

do tego zestawu zagadnień pytania, na ile twórcy testu są w stanie zapanować nad efektami testu, których nie byli w stanie przewidzieć, i czy w drodze nieuniknionej ewolucji, jakiej ulegają wszystkie ważne testy i egzaminy, efekty te mogą zostać wyeliminowane.

Pytania K. Bailey nie obejmują również celów ukrytych testu omawianych niżej przez E. Shohamy. Odpowiedzi na pytania K. Bailey mogą dostarczyć wyłącznie badania empiryczne uwzględniające omawiane przez nią aspekty.

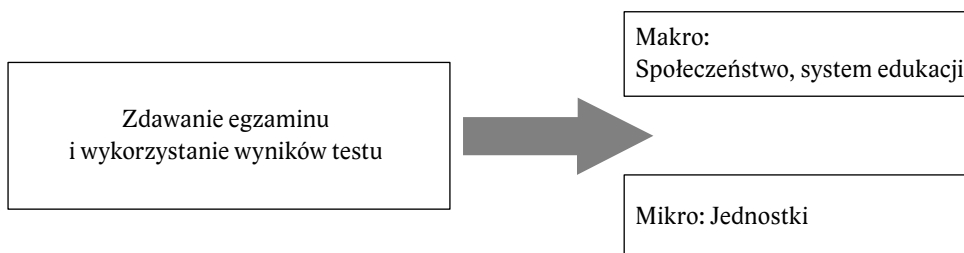
## Zasady Elany Shohamy

Z modelu K. Bailey wynika, że efekt zwrotny jest zjawiskiem niezwykle złożonym. Test wpływa na bezpośrednich uczestników procesu nauczania i uczenia się, tj. nauczycieli i uczniów. K. Bailey rozróżnia efekt zwrotny testu na uczniów i efekt zwrotny testu na program. Swoją własny diagnostyczny model oceniania w zakresie języków obcych w 1992 roku stworzyła E. Shohamy. W jej koncepcji wyeksponowany jest również związek pomiędzy testowaniem i programem. E. Shohamy („Beyond Proficiency Testing” 515) określa sześć zasad, które wyjaśniają choć częściowo działanie efektu zwrotnego. Pierwsza z nich odnosi się do zróżnicowania dwu pojęć osiągnięcia (*achievement*) i biegłości (*proficiency*) – „nie wszystko, czego uczniowie uczą się w szkole (osiągnięcie), jest przydatne w życiu (biegłości)”. Druga zasada dotyczy informacji diagnostycznej i mówi, że efektywne wykorzystanie informacji wypływającej z testu możliwe jest wtedy, gdy taka informacja ma charakter szczegółowy, diagnostyczny, a nie charakter ogólny wyrażony przez ogólny wynik egzaminu. Trzecia zakłada związek nauczania z uczeniem się. E. Shohamy przyjmuje, że „zmiany w systemie będą zachodzić zgodnie z informacją zwrotną wyniesioną z testu”. Czwarta zasada E. Shohamy wprowadza pojęcie sprawców zmiany, którzy muszą być zaangażowani w proces konstruowania testu, jeśli test ma przynieść pozytywny efekt zwrotny. To bowiem oni, nauczyciele i władze administracyjne, są odpowiedzialni za wprowadzanie zmian. Piąta zasada uwzględnia konieczność uzyskania informacji porównywalnej. Programy powinny wykorzystywać informację wypływającą z wyników testu, by ocenić swoją wartość zarówno w odniesieniu do innych programów, jak i zakładanych celów. Ostatnia, szósta zasada głosi potrzebę wprowadzenia testowania komunikacyjnego. Idealny test powinien „odzwierciedlać nowoczesne tendencje w nauczaniu” i koncentrować się na „autentycznych zadaniach i sytuacjach językowych testowanych bezpośrednio”.

Należy jednak zauważyć, że by to sprzężenie zwrotne między testem i programem nauczania przynosiło pozytywne efekty, należy najpierw ustalić ogólne założenia programowe. Kolejność odwrotna, czyli wymuszanie zmian programowych poprzez test, ma zwykle destrukcyjny wpływ na proces uczenia się i nauczania, choćby z powodu braku konsekwencji w przekazywanych treściach. Podstawa programowa narzuca metody nauczania, test może jedynie zweryfikować ich skuteczność. Jasne założenia podstawy programowej definiują cele nauczania i jego plan wynikowy, stosowane metody i techniki wynikają zatem z filozofii edukacyjnej, a nie z konieczności przygotowania do testu. Dlatego konieczna jest spójność między założeniami programowymi i założeniami testu.

## Model Lionela Bachmana i Adriana Palmera

Równoległe do modelu K. Bailey powstał model Lionela Bachmana i Adriana Palmera (Bachman, Palmer).



**Rysunek 2. Model Lionela Bachmana i Adriana Palmera**

Jak wynika z modelu L. Bachmana i A. Palmera, efekt zwrotny testu funkcjonuje na dwóch poziomach: na poziomie makro, obejmującym społeczeństwo i system edukacyjny, oraz na poziomie mikro, obejmującym jednostki. Poprzez jednostki L. Bachman i A. Palmer rozumieją wszystkich, na których testy oddziałują pośrednio lub bezpośrednio, a zatem: zdających, decydentów, przyszłych pracodawców zdających. Kierunek strzałki oznacza, że wartości i cele wpisane w wyniki testu tkwią w społeczeństwach, które z tych wyników robią użytek.

Z tego powodu opinia Johna B. Heatona, że jakość testu przesądza o wywieranym przez niego wpływie na proces uczenia się i nauczania, staje się jeszcze bardziej znacząca. Jeśli bowiem egzamin, zwłaszcza egzamin doniosły może wpływać na cały system



edukacji, tym bardziej jego konstrukcja musi być staranna i przemyślana, sprzężona z celami kształcenia. Gdyby bowiem założyć, że w pewnych przypadkach celem nadrzędnym tego procesu staje się pozytywny wynik testu, mogłoby to budzić najwyższy niepokój. Taka sytuacja prowadzi nieuchronnie do skoncentrowania wysiłków na nauczaniu i uczeniu się pod test, co wywiera zdecydowanie negatywny wpływ na proces edukacji, a w przypadku nietrafnie skonstruowanego testu może się okazać wręcz niszczące. Tak potężny instrument jak test musi być zatem konstruowany i wykorzystywany z wielką rozwagą.

## Typologia zjawiska i jego pochodne

Związki pomiędzy wynikami testów a innymi zjawiskami dydaktycznymi mają wielowymiarowy, złożony, czasem sprzeczny charakter. Co więcej, wpływ tego samego testu aplikowanego różnym osobom w różnym czasie również nie jest jednoznaczny ani jednorodny. Należy zatem dokonać odrębnej analizy poszczególnych elementów tego zjawiska.

## Efekt zwrotny ogólny lub szczegółowy

Efekt zwrotny może mieć charakter ogólny lub szczegółowy. Pod pojęciem efektu zwrotnego ogólnego rozumiemy każdy potencjalny efekt testu (Watanabe). Jeśli na przykład założymy, że test wpływa na sposób pracy nauczyciela, który w przypadku przygotowywania testu staranniej opracowuje zajęcia, oznacza to, że takie zjawisko zachodzi w przypadku różnego typu egzaminów i różnych nauczycieli. Efekt zwrotny szczegółowy odnosi się natomiast do wpływu pewnego aspektu danego testu na proces uczenia się i nauczania. Na przykład wykluczenie danej sprawności z testu może spowodować całkowite lub częściowe wyeliminowanie tejże z procesu nauczania.

## Efekt zwrotny silny lub słaby

Silny efekt zwrotny wpływa na niemal wszystkie zjawiska zachodzące w klasie. Może się zdarzyć tak, że pod wpływem testu nauczyciele przygotowujący do niego uczniów zaczynają uczyć w niemal identyczny sposób, „pod test”. Im silniejszy efekt zwrotny, tym

więcej aspektów procesu nauczania/uczenia się znajduje się w jego zasięgu i pod tym większym wpływem testu znajdują się wszyscy uczestnicy procesu nauczania/uczenia się. Efekt słabszy zachodzi w przypadku odwrotnym, to jest w sytuacji, gdy test wpływa tylko na niektóre (i tylko do pewnego stopnia) aspekty procesu nauczania/uczenia się. Nie ulega jednak najmniejszej wątpliwości, że siła efektu zwrotnego zależy od tego, na ile dany test jest istotny w życiu ucznia (Cheng „Impact”). Dlatego zjawisko efektu zwrotnego analizuje się często w kontekście egzaminów doniosłych.

### **Efekt zwrotny krótkotrwały i długotrwały**

Efekt zwrotny może być zarówno długotrwały, jak krótkotrwały (Cheng, Watanabe, Curtis). Jeśli efekt oddziaływania testu przeciąga się na okres poegzaminacyjny (tj. dalszą karierę edukacyjną uczącego się, np. na okres studiów w przypadku Nowej Matury), można mówić o efekcie zwrotnym długotrwałym. Jeśli na przykład sposób uczenia się, priorytety, nawyki zostałyby zachowane przez uczniów już po rozpoczęciu studiów, oznaczałoby to, że efekt ten jest długotrwały. Jeśli sposób egzaminowania wpływa na uczestników procesu jedynie w okresie przygotowań, jest to efekt krótkotrwały. Ponieważ jednak istnieje ryzyko, że efekt zwrotny może przybrać długotrwały charakter, tym ważniejsza jest taka konstrukcja testu, która wyeliminowałaby jego potencjalnie negatywne skutki na dalszych etapach nauczania.

### **Efekt zwrotny zamierzony i niezamierzony**

Konstruując test, jego autorzy zakładają uzyskanie pewnych celów edukacyjnych. Jak zauważa E. Shohamy („The Power of Tests”), mogą być to cele prezentowane jawnie lub cele ukryte.

Trafność testu można mierzyć właśnie w kategoriach realizacji celów (Messick „Validity”). Na przykład można sobie wyobrazić, że do nowego rodzaju testu wprowadzono testowanie sprawności nie testowanej uprzednio (np. sprawności rozumienia ze słuchu), m.in. po to, by niejako wymusić na nauczycielach i uczących się bardziej intensywną pracę nad tą umiejętnością). Jeśli by się okazało, że istotnie ten sposób testowania wpłynął pozytywnie na umiejętność rozumienia ze słuchu, można by założyć, że cel został osiągnięty. Taki efekt byłby zamierzonym efektem zwrotnym egzaminu. Niemniej jednak najciekawszy dla badaczy pozostaje niezamierzony efekt zwrotny testu pozostający

poza kontrolą autorów – a zatem wszystkie zjawiska występujące w procesie nauczania/uczenia się podczas przygotowań do egzaminu. Ponieważ test wpływa na wszystkich uczestników procesu, jego autorzy zdają sobie sprawę, że test wywiera wpływ na proces nauczania/uczenia się. Tu pole badawcze jest niezwykle szerokie. Interesujący byłby również „bilans zysków i strat”, tzn. ocena, czy zamierzone zmiany pozytywne istotnie przeważają nad stratami wynikającymi z efektów negatywnych.

### Efekt zwrotny uświadomiony i nieświadomiony

Z badań prowadzonych przez piszącą te słowa wynika, że efekt zwrotny egzaminu na proces nauczania lub uczenia się może mieć charakter uświadomiony lub nieświadomiony. Efekt zwrotny uświadomiony to taki efekt zwrotny egzaminu na proces nauczania i uczenia się, z którego zdają sobie sprawę uczestnicy procesu edukacji. Np. wiem, że moim zadaniem na egzaminie będzie napisanie listu formalnego, w związku z czym powtarzam zwroty i kolokacje niezbędne do stworzenia tego rodzaju tekstu i jestem świadoma, że robię to właśnie z uwagi na egzamin. W tym przypadku podanie do wiadomości zdającego formatu wypowiedzi pisemnej skutkuje jego koncentracją na przygotowaniu tej konkretnej formy do egzaminu, z czego on sam zdaje sobie sprawę.

Nieświadomiony efekt zwrotny egzaminu to taki efekt egzaminu, który zachodzi w procesie nauczania, a jednak nie jest wiązany przez uczestnika edukacji z danym testem. Np. wiem, że na egzaminie najwyżej punktowana będzie sprawność czytania, najniżej – znajomość gramatyki, w związku z czym większość czasu przeznaczona na przygotowania poświęcam na czytanie i analizę tekstów, a także rozwiązywanie zadań związanych z ich rozumieniem, rezygnuję natomiast zupełnie z rozwiązywania zadań gramatycznych. Takie priorytety muszą wynikać z przyczyn pragmatycznych tj. chęci uzyskania dobrego wyniku na egzaminie. Odpowiadając zatem na pytanie o typ zadań wykonywanych najczęściej przed egzaminem, wymieniam zadania sprawdzające rozumienie tekstu czytanego i pomijam ćwiczenia gramatyczne. Zapytana jednak w innym pytaniu, czy poświęcam więcej czasu na pewne sprawności w związku z wymaganiami egzaminacyjnymi, odpowiadam przecząco, nie przyznając się do tego, że kieruję się w przygotowaniach względami praktycznymi. Tego rodzaju odpowiedzi mogą mieć czasem związek z efektem Hawthorne, lecz w wielu przypadkach świadczą o tym, że respondenci nie dostrzegają związku pomiędzy swoimi zachowaniami podejmowanymi w ramach przygotowań do egzaminu a samym egzaminem.

Z przeprowadzonych badań wynika, że efekt zwrotny nieświadomiony ma miejsce znacznie częściej w przypadku uczniów niż nauczycieli. Dodać również należy, że efekt zwrotny uświadomiony i nieświadomiony mogą mieć zarówno pozytywny, jak i negatywny charakter.

## Rodzaje trafności testu

Pojęcie trafności testu jest pojęciem złożonym. Jak to ujmuje H. Komorowska, pojęcie trafności testu generalnie wiąże się z pytaniem: „Czy test bada to, co chcemy mierzyć?” Odpowiedź na to pytanie jest tylko pozornie oczywista, ponieważ często test sprawdza mniej umiejętności, niż zakładał autor. Często też sprawdza więcej – np. wykracza poza znajomość języka, jeśli wymaga zapamiętywania szczegółów podczas testu rozumienia ze słuchu lub wbrew intencjom autora sprawdza nie tylko zamierzony element. H. Komorowska podaje tu przykład sprawdzania rozumienia ze słuchu poprzez pisanie streszczeń wysłuchanych tekstów, co bada nie tylko sprawność rozumienia ze słuchu, ale i sprawność pisania w języku obcym. Jeśli taka dodatkowa umiejętność okazuje się dla ucznia trudniejsza niż ta, którą planowaliśmy testować, to na podstawie tak przygotowanego sprawdzianu nie jesteśmy w stanie ustalić stopnia opanowania umiejętności podstawowej i test okazuje się nietrafny.

Zgodnie z badaniami H. Komorowskiej trafność testu można rozpatrywać z różnych punktów widzenia, w zależności od punktu odniesienia.

## Trafność zewnętrzna i wewnętrzna

Według H. Komorowskiej rozróżniamy generalnie dwa rodzaje trafności testu, bez odniesienia do jego efektu zwrotnego. Są nimi trafność zewnętrzna i wewnętrzna. Na pierwszą z nich składają się: trafność prognostyczna (ang. predictive/prognostic validity), czyli zależność pomiędzy wynikami osiąganymi przez ucznia na teście i wynikami, jakie uzyska w nauce w przyszłości, i trafność diagnostyczna lub równoczesna (ang. concurrent validity) ustalana na podstawie kryteriów zewnętrznych, czyli na przykład innego testu na tym samym poziomie trudności zdawanego w tym samym czasie. W przypadku rozbieżności wyników można podejrzewać, iż jeden z testów nie był pod jakimś względem wystarczająco trafny.

W ramach trafności wewnętrznej wyróżniamy: trafność treści (ang. content validity), trafność teoretyczną, nazywaną też konstrukcyjną (ang. construct validity), oraz trafność fasadową, czyli powierzchowną (ang. face validity). Trafność treści można określić jako zgodność treści testu z treściami nauczania, trafność teoretyczna oznacza zgodność sprawdzianu z zakładanymi metodami nauczania, a trafność fasadowa polega na zapewnieniu zdającym formy egzaminu, do której mieliby nastawienie pozytywne, co w praktyce polega na tym, by uczący się przystępujący do testu miał przekonanie, że jest to sprawdzian odpowiedni, sprawiedliwy i ukierunkowany na ważne umiejętności, wokół których koncentrowały się główne aktywności lekcyjne w okresie poprzedzającym test. Według H. Komorowskiej jest to niezwykle ważne dla budowania motywacji i rozwijania pozytywnego stosunku do nauki. Należy zauważyć, że może nastąpić rozdział pomiędzy trafnością fasadową i teoretyczną, jeśli metody i techniki nauczania stosowane przez nauczyciela stoją w sprzeczności z tymi, które zakłada podstawa programowa. Egzamin trafny teoretycznie powinien zabezpieczać przed przypadkowością wyniku, na przykład poprzez odpowiednią liczbę zadań sprawdzających daną kompetencję językową.

## Trafność systemowa

W ujęciu Johna R. Frederiksen i Allana Collinsa występuje dodatkowo pojęcie trafności systemowej (ang. systemic validity). Wiąże się ona z pytaniem, do jakiego stopnia ocena produkcji zdających wpływa pozytywnie na rozwój testowanych sprawności.

J. Frederiksen i A. Collins wychodzą z założenia, że system edukacyjny dopasowuje zarówno program, jak i system nauczania do zakładanych celów, z których jeden to maksymalne poprawianie wyników testów kluczowych. Dlatego tak istotne jest, by egzaminy cechowała trafność systemowa. Systemowo trafny test cechuje bezpośredniość oceny, a także pewien poziom pozytywnie pojmowanego subiektywizmu. Od sprawdzających oczekuje się bowiem zdolności trafnego osądu, analizy i refleksji. J. Frederiksen i A. Collins uważają, że zadania trafne systemowo testu to takie, których wykonanie wymaga szerokiej wiedzy i umiejętności związanych z testowaną dziedziną. Niezmiernie istotna jest również przejrzystość kryterialna, wyrabianie umiejętności samooceny uczących się, a także wyczerpująca informacja zwrotna na temat wyników.

Keith Morrow utrzymuje nawet, że trafność testu należy mierzyć poprzez stopień, do jakiego wpływa on pozytywnie na proces nauczania/uczenia się. Cytowani wyżej J.R. Frederiksen i A. Collins domagają się ponadto, by po kilku latach od wprowadzenia

nowego testu taki stopień zbadać i tym samym zmierzyć trafność wprowadzonego egzaminu. Jak jednak argumentuje S. Messick („Validity and washback”), wyniki tego typu badań mogą być przypadkowe, ponieważ wzrost kompetencji w jakimś konkretnym zakresie może wynikać z innych zmian zachodzących w tym samym czasie w systemie edukacji. Dlatego niekoniecznie należy je postrzegać w kategoriach efektu zwrotnego egzaminu.

Ponadto często zdarza się, że nowo wprowadzony test wpływa na zmianę nauczanych treści, co nie oznacza jednak zmiany sposobu nauczania. Egzamin nierzadko zmienia też zachowania nauczycielskie, pozostając bez wpływu na zachowania uczniowskie. Dlatego według Ch. Aldersona i D. Wall efekt zwrotny jest zjawiskiem na tyle złożonym, że nie można go odnosić bezpośrednio do trafności testu.

Według S. Messicka („Validity and washback” 245) trafność „nie jest cechą testu lub sposobu oceniania, lecz cechą znaczenia jego wyników”. Wnioski płynące z wyników testu należy zatem wyciągać bardzo ostrożnie; szczególna ostrożność jest zalecana zwłaszcza przy podejmowaniu działań. Zwiększenie wysiłków i liczby ćwiczeń, by poprawić kompetencje językowe w ramach danej sprawności, w zakresie której wyniki poprzednich testów wydają się niezadowolające, może wpłynąć pozytywnie na znajomość formy testu oraz strategii egzaminacyjnych. Sam ten fakt nie musi jednak wpływać bezpośrednio na poprawę kompetencji w zakresie danej sprawności. Nie bez znaczenia będzie tu jednak poczucie bezpieczeństwa, komfortu psychicznego, które zawsze stwarza nadzieję na lepszy wynik testu.

S. Messick uważa, że trafność testu to „empiryczna ewaluacja znaczenia i konsekwencji pomiaru z uwzględnieniem czynników zewnętrznych” („Validity and Washback” 241–256). Chodzi o czynniki zewnętrzne, jakie zachodzą w konkretnie badanej sytuacji, czynniki, które mogą spowodować nadinterpretację lub niedocenienie znaczenia wyników egzaminu w danym środowisku. Znaczenie wyniku testu jest bowiem konstrukcją opartą zarówno na pewnych regularnościach wyników oraz ich związkach z innymi zmiennymi, zatem test może być konstrukcyjnie trafny lub nietrafny.

Należy zawsze pamiętać, że przypisywanie pozytywnych zmian zachodzących w procesie nauczania/uczenia się wyłącznie w oparciu o wyniki testów jest pewnym nadużyciem, gdyż zmiany te mogą wynikać z zachowań uczestników procesu nauczania/uczenia się bez związku z samym testem.

O efekcie zwrotnym testu można mówić wyłącznie w sytuacji, gdy zachodzi wyraźny, logiczny związek pomiędzy zmianami w procesie nauczania/uczenia się i cechami testu, które mogą na nie wpływać.

Generalnie negatywny efekt zwrotny testu należy powiązać z wprowadzeniem testów charakteryzujących się niskim stopniem trafności, a pozytywny – z testami o wysokim stopniu trafności.

## Trafność efektu zwrotnego

Pojęcie trafności efektu zwrotnego wprowadza K. Morrow (6), by opisać jakość związku pomiędzy testowaniem i procesem nauczania/uczenia się. Generalnie jest on zdania, że trafność testu powinna być mierzona trafnością efektu zwrotnego, zatem im bardziej pozytywny wpływ dany test wywiera na proces uczenia się i nauczania, tym bardziej jest trafny. Podobnego zdania jest cytowany wyżej C.J. Weir, który proponuje, by ewaluacja testów komunikacyjnych wiązała się ściśle z systematycznym zbieraniem danych nie tylko na temat trafności konstrukcyjnej, trafności treści i trafności powierzchniowej testu, lecz także danych na temat trafności wywołanego przez niego efektu zwrotnego wyrażonej poprzez stopień, w jakim autorom testu udaje się osiągnąć zamierzony efekt zwrotny.

Poglądów K. Morrowa i C.J. Weira nie podzielają jednak Ch. Alderson i D. Wall (116), którzy uważają, że efekt zwrotny to tylko jedna z konsekwencji testu, jaka powinna być brana pod uwagę przy ocenie jego trafności.

Dylan Wiliam (wraz z P. Black) podejmuje natomiast rozumowanie odwrotne niż K. Morrow. Twierdzi, że test nie tyle jest trafny w stopniu, w jakim wywarł pozytywny wpływ na proces nauczania/uczenia się, lecz raczej wywiera pozytywny wpływ w takim stopniu, w jakim jest trafny. Co za tym idzie, test jest trafny, gdy nauczyciele chcą przygotowywać zdających do tego właśnie testu.

Powyższe twierdzenie wydaje się prawdziwe zaledwie w pierwszej części. Trudno bowiem polemizować z poglądem, że to właśnie trafne testy mogą wywrzeć na proces nauczania/uczenia się pozytywny wpływ. Nie znaczy to jednak, że trafny jest taki test, do którego chcą nauczać nauczyciele, ponieważ mogą się oni kierować indywidualnymi, niekoniecznie obiektywnie słusznymi preferencjami co do wyboru metody. Jeśli są bowiem przykładowo zwolennikami metody kognitywnej, test o charakterze wyłącznie komunikacyjnym może nie zostać przez nich zaakceptowany. Z kolei pogląd, że trafność testu należy mierzyć stopniem, w jakim wywarł on pozytywny wpływ na proces uczenia się/nauczania, wydaje się nazbyt ogólny – naszym zdaniem należy bowiem przyjąć, że żaden test nie wywiera jednoznacznie pozytywnego (ani negatywnego) wpływu na proces

edukacji i jego uczestników. Charakter tego wpływu zależy jest od wielu czynników, m.in. od poziomu wyjściowych kompetencji językowych uczących się w momencie rozpoczęcia przygotowań do testu, celu testu, jego konsekwencji i wielu innych. Dlatego ten sam test może wywierać jednocześnie efekt zwrotny zarówno pozytywny, jak i negatywny, zarówno na uczniów, jak i nauczycieli.

### **Efekt zwrotny pozytywny i negatywny**

Jak wspomniano wyżej testy wywierają na uczestników procesu nauczania/uczenia się wpływ zarówno pozytywny, jak i negatywny; wpływy te jednak nie zawsze można określić jednoznacznie, ponieważ warunkują je zarówno czynniki związane z samym testem, jak też czynniki zewnętrzne. Dlatego pojęcia efektu zwrotnego pozytywnego i negatywnego są w pewnym stopniu względne. Niemniej jednak można analizować pod tym kątem wpływ egzaminów na zdających, zwłaszcza na etapie pilotażowym. Jest to zresztą jeden z postulatów Thomasa Kellaghana i Vincenta Greaney, którzy sądzą, że należy dokonać prognostycznej ewaluacji trafności egzaminów zewnętrznych, by ocenić, czy istotnie realizują stawiane przed nimi cele.

Należy zauważyć, że w pierwszej fazie po implementacji testy zwykle ulegają drobnym modyfikacjom. Przyczyną tych modyfikacji jest wystąpienie czynnika niezamierzonego negatywnego efektu zwrotnego, bywa nim również konieczność weryfikacji kryteriów oceny, które nie przystają do obserwowanej rzeczywistości (np. zbyt wygórowane lub zaniżone wymagania, brak jasności, niespójność etc). Dlatego przy badaniach efektu zwrotnego należy wziąć pod uwagę również ten czynnik. Ważna jest też analiza wyników wprowadzonych już testów, konsultacje z uczestnikami procesu edukacji na każdym etapie wprowadzania reform i dokonywanie zmian na podstawie wniosków płynących z podejmowanych czynności.

### **Efekt zwrotny pozytywny**

Trudno sobie wyobrazić, aby zamiarem autorów testów było wywołanie skutków negatywnych, zatem zwykle przyjmuje się, że efekt zwrotny pozytywny jest zamierzony, a efekt zwrotny negatywny – niezamierzony. Należy tu jednak zastrzec, że sam pomysł, by wprowadzać zmiany programowe poprzez zmiany w sposobie testowania wydaje się



bardzo kontrowersyjny. Jeśli bowiem nawet skuteczność takiego działania jest wysoka, to naszym zdaniem chronologia działań powinna być odwrotna: zmiany w testowaniu powinny wynikać z przemyślanych zmian programowych. Niemniej jednak, jakość efektu zwrotnego nie zawsze jest jednakowa dla wszystkich uczestników procesu. Ten sam test może oddziaływać pozytywnie na jednych i negatywnie na drugich. Nie decyduje o tym zresztą sama formuła testu, ale również wiele czynników zewnętrznych niezależnych od autorów testu, takich np. jak liczebność nauczanych grup, warunki nauczania, cele jednostkowe w społeczności zdającej ten sam test, przygotowanie nauczyciela do zawodu, jakość materiałów dydaktycznych i wiele innych. Ważnym czynnikiem jest również to, czy autorzy testu i uczestnicy procesu uznają te same osiągnięcia za istotne, znaczące dla rozwoju językowego. W sytuacji, w której nowy test jest oficjalnie narzędziem zmiany polityki językowej, jego trafność może być rozpatrywana właśnie poprzez pryzmat pozytywnego, zamierzonego efektu zwrotnego. Strategię wykorzystywania testów streszczają Lewis Elton i Diana Laurillard, twierdząc, iż jest to najszybszy sposób, by zmienić sposób uczenia się. Według D. Wall („Test Impact”) musi upłynąć kilka lat, by ocenić zamierzony, pozytywny efekt zwrotny nowego testu.

Wydaje się, że pozytywny efekt zwrotny łatwiej jest osiągnąć, jeśli uczestnicy procesu są pozytywnie nastawieni do zmian i chcą osiągnąć zamierzone cele. Pozytywne oddziaływanie testu na uczestników procesu jest jednak sprawą względną. Według Stephena P. Heynemana na przykład wielu zwolenników oceniania poprzez testy uważa za zaletę umiejętność poddania się „tresurze”. Innym argumentem jest ten podawany przez Iana Pearsona, że rozwiązywanie testów próbnych może stać się korzystną dla uczących się aktywnością lekcyjną. I. Pearson dostrzega jednak drugą stronę zagadnienia, pisząc, że zasada ta działa również w przeciwną stronę, gdyż interesujące aktywności lekcyjne mogą być stosowane w testowaniu. A. Davies („Follow my leader”) wskazywał na pozytywną rolę testu pojmowanego jako prekursora zmian, który może w rezultacie doprowadzić do twórczych zmian programowych.

Badając pozytywny, zamierzony efekt testowania na proces nauczania/uczenia się nie można jednak abstrahować od założeń programowych – zwłaszcza, jeśli test stał się ich katalizatorem. Jedynie porównując założenia programowe z efektem zwrotnym testu, można sprawdzić, czy test spełnił zamierzone zadanie.

Jak wspomiano wcześniej, aby wywołać dobroczynny, pozytywny efekt testu, zarówno nauczyciele, jak i uczniowie muszą rozumieć dokładnie jego cel (Hughes „Testing”). Muszą również uznawać jego wyniki za wiarygodne i sprawiedliwe (Shohamy „Beyond Proficiency Testing”). Pozytywny wpływ efektu zwrotnego zaznacza się

wówczas, gdy test sprawdza treści zawarte w programie (Bailey). Tylko wtedy można podchodzić poważnie zarówno do wyników pozytywnych egzaminu, jak i ewentualnych niepowodzeń uczniów. Istotne jest także, by test znajdował oparcie w solidnych podwalinach teoretycznych nowoczesnego nauczania i jasno sformułowanych celach nauczania.

Równie istotnym zagadnieniem jest autentyczność zadań. Wówczas przygotowujący się do testu ćwiczą działania językowe, jakie przydadzą im się w życiu. Wspomina o tym zarówno Majorie Wesche, jak i G. Buck. „Naturalność, autentyczność, atrakcyjność i duże prawdopodobieństwo napotkania tego rodzaju sytuacji w życiu to podstawowe cechy dobrze prowadzonej kontroli i oceny” – pisze H. Komorowska (30). Dodaje dalej, że skoro uczniowie zawsze uczą się „pod test”, warto sprawdzać to, co naprawdę istotne, zatem „praktyczność testu to także autentyczność komunikacyjna zadań” (ibid.). A. Hughes („Testing”) doradza stosowanie testowania bezpośredniego i testowania umiejętności, jakie chce się promować.

Według Petera Doye’a optymalnym rozwiązaniem byłoby przystawanie zadań testowych do sytuacji występujących w prawdziwym życiu. Nie jest to jednak zapewne możliwe do osiągnięcia, ponieważ często nawet najbardziej starannie przygotowanym zadaniom testowym towarzyszy brak naturalności. Dlatego test nie może stanowić w pełni udanej i wiernej reprodukcji sytuacji występujących w życiu. Można jednak dążyć do zbliżenia zadań egzaminacyjnych do sytuacji występujących w życiu, w nadziei na to, że wyuczone zachowania zostaną przywołane w najbardziej stosownym momencie. Już w 1980 roku Brendan J. Carroll (11) wspominał o tym, że zadania testowe powinny mieć cechy naturalnej komunikacji dnia codziennego w przeciwieństwie do sztucznych sytuacji bodźca i reakcji. B. Carroll podkreślał konieczność takiej konstrukcji zadań, by wymagały one działań interaktywnych. Według B. Carrolla kontekst wymiany myśli powinien być jak najbardziej realistyczny, wyzuty z uproszczeń, szkolnych zwrotów. Przeciwnie, taki egzaminacyjny dyskurs nie powinien się różnić od codziennych rozmów ze wszystkimi ich cechami naturalnymi, takimi na przykład jak odgłosy w tle. K. Morrow jest zdania, że z tej perspektywy równie ważna jest autentyczność materiałów egzaminacyjnych, jak i oczekiwanych wypowiedzi (Doye).

Pogląd K. Morrowa podziela S. Messick („Validity and washback”), który uważa, że zadania testowe powinny mieć charakter autentycznych działań językowych, jakie można napotkać w życiu codziennym.

Są powody, by sądzić, że kolejnym czynnikiem promującym dobroczynny efekt zwrotny testu jest dążenie do jak największej autonomii uczących się. Umiejętność samooceny – jak podsumowuje Tibor von Elek (60) – sprawia, że uczący się przyjmują

większą odpowiedzialność za ocenę swoich umiejętności i postępów, potrafią określić swoje słabe strony, biegłość językową w danym momencie i zdać sobie sprawę z różnicy, jaka dzieli tę biegłość od zamierzonej. Dzięki autonomii uczący się zyskują większą motywację do osiągnięcia celu i dlatego należy ich bardziej zaangażować w proces oceniania (Alderson).

Według S. Messicka, aby uzyskać optymalny efekt zwrotny, różnica pomiędzy aktywnościami typowymi dla codziennej nauki języka i tymi, których wymaga przygotowanie do testu, powinna być jak najmniejsza („Validity and washback” 241–242). Podkreśla on jednak, że zjawisko efektu zwrotnego jest zbyt skomplikowane, by można było jego analizę ograniczać do tak prostych zależności, jak sugerował to cytowany wcześniej J. Heaton. W odbiorze społecznym bowiem nawet słaby test może być odbierany pozytywnie, dobry – negatywnie. Dlatego Ch. Alderson i D. Wall utrzymywali, że zjawisko efektu zwrotnego nie musi być związane z trafnością testu. Takiego poglądu nie podziela Robert Linn (16), który jest zdania, że aby właściwie ocenić trafność testu, należy dokonać ewaluacji efektu zwrotnego, jaki ów test wywiera na proces nauczania i uczenia się.

Jednym z warunków koniecznych do uzyskania pozytywnego efektu zwrotnego testowania jest zapewnienie zdającym dostępu do szczegółowych sprawozdań z wyników testów, obszerniejszych niż te, które stosuje się zwykle po przeprowadzeniu egzaminów (Bailey). Standardowa informacja zwrotna wyrażona poprzez ocenę lub procenty nie wystarcza, ponieważ nie przynosi szczegółowej informacji na temat kompetencji językowych zdających w obrębie poszczególnych sprawności. Dlatego informacja zwrotna powinna zawierać dokładny opis osiągnięć uczących się, zwłaszcza że najczęściej kompetencje językowe zdających różnią się mocno w ramach poszczególnych sprawności, i przedstawienie wyniku w formie ogólnej, niezależnie od tego, czy jest to ocena tradycyjna, czy procentowa, nie wystarczy do wyciągnięcia odpowiednich wniosków na temat dalszego procesu uczenia się i nauczania. Pozytywny efekt zwrotny długotrwały wymaga zatem stosownej formuły opisu wyników.

Według D. Wall („The impact of high-stakes examinations”) łatwiej jest osiągnąć pozytywny efekt zwrotny testu, jeśli cele i treści testowane na egzaminie odpowiadają celom i treściom podstawy programowej. Należy tu jednak zastrzec, że powyższe stwierdzenie jest prawdziwe jedynie wówczas, gdy owa podstawa programowa jest przemyślana i stawia właściwe cele edukacyjne.

D. Wall uważa, że pozytywny efekt zwrotny może objąć trzy aspekty nauczania: jego treści, metody i sposób oceniania. Jeśli chodzi o efekt zwrotny na treści nauczania, to nauczyciele zwracaliby w tym przypadku uwagę na wszystkie części podręcznika, ponieważ

zdawaliby sobie sprawę, że egzamin może zawierać najróżniejsze typy zadań. To samo odnosiłoby się do nauczanych sprawności, które miałyby ten sam ciężar wagowy w podręczniku, co na egzaminie. Pozytywnym efektem zwrotnym w zakresie metod nauczania byłoby wykorzystywanie metod wynikających z założeń programowych, odzwierciedlonych jednocześnie w podręczniku i opisanych w książce nauczyciela. W zakresie oceniania pozytywnym efektem zwrotnym byłyby nauczycielskie testy cząstkowe sprawdzające opanowanie materiału zawartego w podręcznikach, gdyż na tej właśnie zasadzie konstruowano egzaminy. Takie podejście do zagadnienia zakłada interakcję między testem i uczestnikami procesu nauczania.

Interakcje między wszystkimi uczestnikami procesu edukacji są kluczem do tego, by egzaminy stanowiły element konstruktywny procesu nauczania, a nie – jak to się czasem dzieje – czynnik destrukcyjny. Takie interakcje powinny zachodzić na kilku poziomach. Poziom pierwszy to relacja pomiędzy władzami edukacyjnymi odpowiedzialnymi za wprowadzenie testu a nauczającymi. Gdyby nauczający odgrywali czynną rolę w tworzeniu testów, wręcz je współtworzyli, mogłoby to zniwelować szkodliwy podział na „my” – szkoła (nauczyciele oraz uczniowie) i „oni” (autorzy testów oraz władze edukacyjne). W tej sytuacji nauczający nie musieliby stosować się do narzuconych formatów, lecz odgrywaliby twórczą rolę w kształtowaniu procesu edukacji. Co więcej, udział nauczycieli w tworzeniu zadań egzaminacyjnych gwarantowałby, że stanowią one rzeczywiste odbicie narzucanych treści, a nie odwrotnie, jak to się dzieje w przypadku, gdy formuła testu ogranicza treści programowe. Zadania egzaminacyjne byłyby wówczas bardziej różnorodne, co zniwelowałoby przynajmniej w pewnym stopniu problem nauczania „pod test”. Współtworzenie testów mogłoby również sprawić, że nauczyciele czuliby się bardziej odpowiedzialni za proces nauczania.

Niezmiernie ważne są również konsultacje społeczne, wsłuchanie się w głos wszystkich uczestników procesu edukacji poprzez nieustanną ewaluację wprowadzanych egzaminów i ich modyfikacje. Opinia egzaminowanych pozwala na spojrzenie na wprowadzany egzamin z innej perspektywy. Niejednokrotnie dzięki takim opiniom można by odkryć niezamierzone efekty wprowadzanych testów, opinie uczących się pozwalała by również zrozumieć, jakie są ich nadzieje związane z nauką oraz oczekiwania w stosunku do otrzymywanej informacji zwrotnej.

Kolejną ważną interakcją w przypadku egzaminów zastępujących egzaminy wstępne na wyższe uczelnie jest relacja pomiędzy ich twórcami i reprezentantami uczelni wyższych, którzy mogliby formułować swoje oczekiwania w stosunku do kandydatów, wpływając pozytywnie na treści testowane przez egzamin. Cenny byłby również kontakt

pomiędzy nauczającymi i uczniami a twórcami podręczników, ponieważ pozwoliłoby to na poprawienie jakości materiałów dydaktycznych. Wymaga to jednak wzmoczonego wysiłku koordynatorów procesu edukacyjnego oraz powołanie zespołów roboczych odpowiedzialnych za jego jakość.

## Efekt zwrotny negatywny

Badaczy interesuje jednak w równym stopniu, a czasem nawet bardziej, negatywny efekt zwrotny testu na proces nauczania/uczenia się, który zawsze stanowi potencjalny problem. Jeszcze w czasach, gdy nie istniało nawet pojęcie efektu zwrotnego, choć samo zjawisko występowało zawsze, autorzy testów i dydaktycy wiedzieli, że test wywiera duży wpływ zarówno na uczniów, jak i nauczycieli, którzy pod jego wpływem zmieniają często strategie, cele, techniki i całą filozofię NUJ2.

Już w latach pięćdziesiątych stwierdzono (Vernon), że nauczyciele przygotowujący uczniów do ważnych egzaminów pomijają często pewne treści programowe, które nie są istotne w tym konkretnym teście. Strategie testowania stawały się zatem strategiami nauczania (Davies „Oral English”). Stephen Wiseman krytykował zajęcia poświęcone całkowicie przygotowaniu do testu, ponieważ uznawał, że uczą one bardziej strategii egzaminacyjnych niż języka.

Jak zauważyli Ch. Alderson i D. Wall, niewłaściwie skonstruowany test musi przynieść negatywny efekt zwrotny, nawet jeśli proces nauczania nie budzi zastrzeżeń. Taka sytuacja wynika z faktu, że nauczyciele uczą „pod test”, do testu przygotowują się uczniowie, a zatem zadania testowe będą wyznaczać zarówno strategie nauczania, jak i uczenia się.

K. Bailey jest zdania, że przygotowujący się do testu ćwiczą zwykle zadania podobne w formie do zadań testowych, czytają dużo w języku docelowym, uczą się słownictwa i reguł gramatycznych, uczestniczą w interakcjach w języku docelowym (np. prowadzą konwersacje w tym języku, słuchają audycji radiowych w języku, którego się uczą, oglądają obcojęzyczne programy telewizyjne, stosują strategie, jakich wymagają od nich zadania testowe, zapisują się na kursy językowe przygotowujące do egzaminu, proszą o wsparcie i komentarze na temat podejmowanych przez nich działań językowych, zapisują się na zajęcia uzupełniające standardowy proces nauczania lub wręcz się ich domagają, a równocześnie – co wydaje się wręcz absurdalne – opuszczają zajęcia, by w domu przygotowywać się do testu.

Należy zauważyć, że efekt zwrotny negatywny jest pojęciem w pewnym stopniu względnym (coś, co szkodzi jednemu, może pomóc innym), choćby niepokój, jaki odczuwają zdający przed testem, może być przez jednych odbierany jako czynnik mobilizujący, pozytywny, i jako bardzo negatywny przez innych.

Stan emocjonalny uczącego może zarówno hamować rozwój jego kompetencji językowych, jak je wspierać. Mowa tu o negatywnym efekcie zwrotnym, jaki test wywiera na ucznia. Nauczyciele odpowiedzialni za wynik testów mają natomiast wpływ na promowanie takich zachowań uczniowskich, jakie uważają za właściwe w procesie przygotowań, toteż można założyć, że w jakimś stopniu je inspirują.

Co więcej, jak zauważa Qi Luxia, trudno jednoznacznie stwierdzić, czy efekt zwrotny danego testu jest negatywny bez odniesienia tego testu do programu nauczania. Dopiero po dokonaniu dokładnej charakterystyki założeń programowych można tego typu wnioski wyciągać. Jednoznaczne wnioski wysnuć będzie jednak jeszcze trudniej, jeśli sam program budzi wątpliwości. Zatem jeżeli test jest narzędziem kontroli programu przed wprowadzeniem nowego testu, należy dokładnie określić zakładane zmiany, jakie mają w programie nastąpić, ponieważ bez tego punktu odniesienia taka ocena nie jest możliwa. Wówczas można jedynie analizować zjawiska, jakie zachodzą w procesie edukacyjnym po wprowadzeniu danego testu, pamiętając przy tym, że nie wszystko, co dzieje się w klasie, wynika z procedur egzaminacyjnych. Część zjawisk zarówno negatywnych, jak i pozytywnych może być następstwem innych okoliczności.

Generalnie można zatem założyć, że źle skonstruowany test nie może przynieść pozytywnego efektu zwrotnego. Nie znaczy to jednak, że jego gwarantem jest test skonstruowany poprawnie.

E. Shohamy („Beyond Proficiency Testing”) definiuje okoliczności, które mogą doprowadzić do negatywnego efektu zwrotnego. Podkreśla, że niekorzystne jest ciągłe skupianie uwagi na biegłości językowej uczniów zamiast na działaniach, które tę biegłość kształtują. Równie niebezpieczna jest sytuacja, kiedy testy tworzone są w sposób autorytarny bez udziału tych, którzy są odpowiedzialni za wprowadzanie zmian, tj. nauczycieli. Testy nie mogą wpłynąć pozytywnie na proces nauczania, jeśli ich wyniki nie są poddawane szczegółowej analizie, która z kolei może być przyczyną wprowadzania kolejnych zmian.

D. Wall („Test Impact”), która określiła warunki sprzyjające wystąpieniu pozytywnego efektu zwrotnego, zdefiniowała również sytuację, w której efekt zwrotny testu mógłby być negatywny. Jej zdaniem działałoby się tak, gdyby realizowany podręcznik i egzamin miały inne cele, ponieważ nauczyciele wykazywaliby wówczas tendencję, by uczyć

„pod egzamin”. W zakresie treści nauczania nie przywiązywaliby wagi do wszystkich części podręcznika, ponieważ zdawaliby sobie sprawę, że pewne umiejętności nie są oceniane na egzaminie i dlatego lepiej koncentrować się na tych, które podlegają ocenie. A jeśliby nawet nauczyciele danych sprawności nauczali, mogliby celowo pomijać aktywności nieobecne na teście, gdyż szkoda by im było na nie czasu.

Kolejnym negatywnym efektem zwrotnym byłyby tu rezygnacja z podręcznika na rzecz materiałów powielających treści oceniane na egzaminie, takich jak egzaminy próbne lub inne materiały prezentujące przykłady zadań egzaminacyjnych.

Należy jednak zauważyć, że D. Wall zakłada sytuację, w której podręcznik jako taki nie jest docelowo pisany pod kątem egzaminu. Zdarza się jednak, że materiały dydaktyczne, jakie powstają po wprowadzeniu nowego egzaminu, są konstruowane w taki sposób, by ilustrowały jak najdokładniej zadania egzaminacyjne. Co gorsza, jak wynika ze wspomnianych wcześniej badań właśnie tak zaplanowane podręczniki cieszą się uznaniem wśród uczniów i nauczycieli.

Negatywny efekt zwrotny testu w zakresie metody nauczania występuje według D. Wall w sytuacji, w której nauczyciele są skłonni stosować tylko takie metody nauczania, które uznają za bardziej pomocne czy wręcz wydajne w procesie przygotowywania do konkretnego egzaminu. Niektórzy byłiby nawet gotowi poświęcić skuteczne metody stosowane wcześniej, gdyby się okazało, że w danym przypadku nie przyniosą efektów.

W zakresie kryteriów oceniania nauczyciele konstruowaliby testy cząstkowe stanowiące odbicie raczej poprzednich egzaminów niż treści podręcznika. Zadania testowe stanowiłyby w zasadzie modyfikację zadań testowych z poprzednich egzaminów lub byłyby to wykorzystane wcześniej zadania. Kryteria oceny byłyby również takie same jak te, które zostały przyjęte przez twórców egzaminu.

Należy stwierdzić, że negatywny efekt zwrotny egzaminów występuje zawsze, ważny jest jednak w jego ocenie bilans zysków i strat. William J. Popham (679–682) twierdzi, że jedynym dowodem na pozytywny wpływ egzaminów na system edukacji jest procentowy wzrost liczby uczniów, którym udało się opanować umiejętności podstawowe. Jeśli ma rację, to nie brak powodów do pesymizmu, tym bardziej, że nie dowiedziono, by nawet prawidłowe testowanie przyczyniało się do rozwoju umiejętności bardziej złożonych i poznawczych. Taki stan rzeczy musi bowiem budzić obawy o to, że w sformalizowanym gąszczu nie do końca jasnych standardów i nie do końca sprecyzowanych kryteriów zginą indywidualności, których cechuje oryginalność myślenia i pokusa poszukiwania niecodziennych rozwiązań. Fakt, iż większy procent uczniów może dzięki trafnym testom opanować umiejętności podstawowe, stanowi tylko pewne pocieszenie,

z całą pewnością nie rekompensatę, ponieważ to wybitne i twórcze jednostki gwarantują rozwój nauki, gospodarki oraz innych dziedzin życia. Zwiększona produkcja „wykonawców” nie przysporzy twórców. Dlatego należy dołożyć wszelkich starań, by tak modna standaryzacja nie zabijała indywidualności, co zdarza się niestety zbyt często, również w polskiej szkole.

## Podsumowanie

Efekt zwrotny egzaminu rzadko bywa jednoznacznie negatywny lub jednoznacznie pozytywny – jest wszak zależny od badanego uczestnika i elementu procesu edukacji. Niezwykle ważny jest bilans efektu zwrotnego egzaminu oraz analiza zjawiska w odniesieniu do poszczególnych uczestników procesu uczenia się i nauczania, samego procesu i jego produktów, czyli materiałów nauczania i programów, a także modyfikowanych wersji testów i elementu procesu edukacji.

Egzamin może wywierać inny efekt zwrotny na uczniów niż na nauczycieli, przy czym jego oddziaływanie nie musi być jednorodne w całej grupie badanych. Wydaje się bowiem, że w przypadku efektu zwrotnego egzaminu na ucznia wiele zależy m.in. od takich czynników, jak kompetencje językowe ucznia (można przyjąć, że jakikolwiek obowiązkowy egzamin zmusza go do pracy), jego stosunek do obowiązków szkolnych (czy uczy się wyłącznie pod presją egzaminu), plany na przyszłość (jaką rolę odgrywa dany egzamin z perspektywy jego planów edukacyjnych i zawodowych). Efekt zwrotny egzaminu na ucznia może być silny, słaby, krótkotrwały lub długotrwały, uświadomiony bądź nieświadomiony. Jego charakter wymaga zawsze szczegółowych badań. Należy też pamiętać o czynnikach zewnętrznych, takich jak typ szkoły, do której uczęszcza uczeń, miejsce zamieszkania czy jego sytuacja finansowa, które mogą wpłynąć na dostęp do materiałów dydaktycznych i dodatkowych form kształcenia, co z kolei determinuje proces uczenia się.

W przypadku efektu zwrotnego egzaminów na nauczycieli sprawa przedstawia się podobnie – efekt zwrotny na nauczycieli może mieć różnoraki charakter w zależności od czynników takich jak doświadczenie zawodowe, przyjmowane koncepcje nauczania, styl dydaktyczny, warunki nauczania oraz wiele innych zmiennych. Ten sam egzamin przyjmowany bywa entuzjastycznie przez jednych i odrzucany przez drugich, choćby ze względu na różne oczekiwania wobec testu stanowiącego zawsze odzwierciedlenie pewnej koncepcji nauczania, którą niekoniecznie muszą podzielać. Należy również pa-



miętać, że wyniki testów, jakie osiągają ich uczniowie, mogą wpływać na ich karierę zawodową nauczycieli, co z kolei determinuje często ich stosunek do samego egzaminu.

Ważnym elementem jest efekt zwrotny testu na materiały, programy nauczania i same testy, które zwykle podlegają ewolucji. Testy wpływają bowiem w sposób zasadniczy na konstrukcję materiałów nauczania i programów, co zwykle skutkuje nauczaniem „pod test” wpływającym niejednokrotnie na negatywną lub pozytywną ocenę samego egzaminu. Widać zatem, iż w przypadku efektu zwrotnego niezwykle interesujące są właśnie wpływy wzajemne egzaminów i szeroko pojętego procesu edukacji (z uwzględnieniem jego uczestników, procesów i produktów).

Podsumowując, należy podkreślić, że efekt zwrotny jest zależny od typu testu, a zwłaszcza jego znaczenia dla kariery edukacyjnej ucznia i sposobu organizacji egzaminu. Inny jest bowiem efekt zwrotny sprawdzianów o charakterze wewnętrznym (np. zwykłej klasówki czy sprawdzianu) niż egzaminu mającego znaczący wpływ na losy ucznia. Takie znaczące egzaminy mogą mieć zarówno charakter wewnętrzny, jak i zewnętrzny. Naszym zdaniem sposób organizacji tych egzaminów stanowi jednak osobny element decydujący o charakterze efektu zwrotnego i jego sile, co jednak stanowi już przedmiot odrębnych rozważań.

## Bibliografia

- Alderson, J.C. “Learner-centered testing through computers international issues on individual assessment.” J.H.A.L. de Jong. i D.K. Stevenson, ed. *Individualizing the assessment of language abilities*. Clevedon: Multilingual Matters, 1990. 20–27.
- Alderson, J.C., Hamp-Lyons, L. “TOEFL preparation courses: A study of washback.” *Language Testing* 13 (1996): 280–297.
- Alderson, J.C., Wall, D. “Does washback exist?” *Applied Linguistics* 14 (1993): 115–129.
- Bachman, L. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press, 1990.
- Bachman, L.F., Palmer, A.S. *Language Testing in Practice*. Oxford: Oxford University Press, 1996.
- Bailey, K. “Working for washback: A review of the washback concept in language testing.” *Language Testing*, 13 3 (1996): 257–279.
- Biggs, J.B. “The psychology of assessment and the Hong Kong scene.” *Bulletin of the Hong Kong Psychological Society* 29 (1992): 1–21.

- . "Assumptions underlying new approaches to educational assessment." *Curriculum Forum*, 4 2 (1995): 1–22.
- Buck, G. "Testing listening comprehension in Japanese university entrance examinations." *JALT Journal* 10 (1988): 15–42.
- Carroll, B.J. *Testing communicative* Oxford: Pergamon Press, 1980.
- Cheng, L. "How does washback influence teaching? Implications for Hong Kong." *Language and Education* 11 (1997): 38–54.
- . "Impact of a public English examination on students' perceptions and attitudes toward their English language learning." *Studies in Educational Evaluation* 24 (1998): 279–301.
- Cheng, L., Watanabe, Y. i Curtis, A. *Washback in language testing. Research contexts and methods*. Mahwah, New Jersey, London: Lawrence Erlbaum Associates Publishers, 2004.
- Davies, A. "Oral English testing in West Africa." A. Davies, ed. *Language testing symposium: A psycholinguistic approach*. London: Oxford University Press, 1968. 151–179.
- . "The construction of language tests." J.P.B. Allen, A. Davies, eds. *Testing and Experimental Methods*, vol. 4, *The Edinburgh Course in Applied Linguistics*. London: Oxford University Press, 1977. 38–104.
- . "Follow my leader: is that what language tests do?" Y.P. Lee, A.C.Y. Fok, R. Lord i G. Low, eds. *New Directions in Language Testing*. Oxford: Pergamon Press, 1985. 3–13.
- Doye, P. (1991). "Authenticity in foreign language testing." S. Anivan, ed. *Current developments in language testing*. Singapore: Regional Language Centre, 1991. 103–105.
- von Elek, T. "A test of Swedish as a second language; an experiment in self-assessment." Y.P. Le, C.Y.Y. Fok, R. Lord, G. Low, eds. *New dimensions in language testing*. Oxford: Pergamon Press, 1985. 47–58.
- Elton, L., Laurillard, D. "Trends in student learning." *Studies in Higher Education* 4 (1979): 87–102.
- Finocchiaro, M., Sako, S. *Foreign language testing; A practical approach*. New York: Regents Publishing Co., 1983.
- Frederiksen, N. "The real test bias: Influence of tests on teaching and learning." *American Psychologist* 39 (1984): 193–202.
- Frederiksen, J.R. i Collins A. (1989). "A systems approach to educational testing." *Educational Researcher* 18 9 (1989): 27–32.
- Gates, S. "Exploiting washback from standardized tests." J.D. Brown, S.O. Yamashita, eds. *Language Testing in Japan*. Tokyo: Japanese Association for Language Teaching, 1995. 101–106.
- Heaton, J. B. *Classroom testing*. Harlow: Longman, 1990.

- Heyneman, S.P. "Uses of examinations in developing countries." Selection, research, and education section management. *International Journal of Educational Department*, 74 (1987): 251–253.
- Hughes, A. *Testing for language teachers*. Cambridge: Cambridge University Press, 1989.
- Hughes, A. *Backwash and TOEFL 2000*. Tekst nieopublikowany, Educational Testing Service, 1994.
- Kellaghan, T., Greaney, V. *Using examinations to improve education: a study of fourteen African countries*. Washington, DC: The World Bank, 1992.
- Komorowska, H. *Sprawdzanie umiejętności w nauce języka obcego. Kontrola – ocena – testowanie*. Warszawa: Fraszka Edukacyjna, 2002.
- Lewkowicz, J., Zawadowska-Kittel, E. "Impact of the new school leaving examination in Poland." *Research Notes*, 34, Cambridge: ESOL, 2008.
- . "Classroom-based assessment – the Polish experience." D. Tsagari, I. Csepes, eds. *Classroom based language assessment*. Frankfurt am Mein, Berlin, Bern, Brussels, New York, Oxford, Wien: Peter Lang, 2011. 147–167.
- Li, X., Gui S., Li. W. "The design of the NMET and ELT in middle schools." *English Language Teaching and Research in Primary Schools and Middle Schools* 1 (1990): 1–27.
- Linn, R.L. "Evaluating the validity of assessments. The consequence of use." *Educational Measurement; Issues and Practice* 16 2 (1997): 14–16.
- Longman dictionary of contemporary English* Pearson Education Ltd. Essex: Harlow, 2003.
- Luxia, Q. *The intended washback effect of the national matriculation English test in China tentions and reality*. Beijing: Foreign Language Teaching and Research Press, 2004.
- Madsen, H. *Techniques in testing*. Oxford: Oxford University Press, 1983.
- Mehrens, W.A., Kaminsky, J. "Methods for improving standardized test scores: fruitful, fruitless or fraudulent?" *Educational Measurement: Issues and Practices* 8 1 (1989): 14–22.
- Messick, S. "Validity." R. Linn, ed. *Educational measurement* (3rd ed.) New York: Macmillan, 1989: 13–103.
- . "Validity and washback in language testing." *Language Testing* 13 3 (1996), 241–156.
- Morrow, K. "The evaluation of tests of communicative performance." M. Portal, ed. *Innovations in Language Testing*. Windsor: NFER/Nelson, 1986. 1–13.
- New Webster Comprehensive Dictionary*. Trident Press International, Com edition (2003).
- Niemierko, B. *Ocenianie szkolne bez tajemnic*. Warszawa: WSIP, 2002.
- Pearson, I. "Tests as levers of change." D. Chamberlain, R.J. Baumgardner, ed. *ESP in the classroom: Practice and evaluation*. London: Modern English Publications in association with the British Council, 1988. 98–107.

- Popham, W.J. "The merits of measurement-driven instruction." *Phi Delta Kappan* May (1987): 679–682.
- Poszytek, P. *European standards of testing language proficiency and the Polish "Nowa Matura" examinations in English*. Kraków: Universitas, 2008.
- Shawcross, P. *What do we mean by washback effect of testing?* JALT. Aviation English Services, 2002.
- Shohamy, E. "Beyond proficiency testing: A diagnostic feedback testing model for assessing foreign language learning." *The Modern Language Journal* 76 4 (1992): 513–521.
- . *The power of tests: A critical perspective on the use of language tests*. London: Pearson Education, 2001.
- Shohamy, E., Donitsa-Schmidt, S., Ferman, I. "Test impact revisited: Washback effect over time." *Language Testing* 13 3 (1996): 299–317.
- Słownik angielsko-polski Exeter*. Warszawa: Kastor, 2006.
- Spolsky, B. "The examination-classroom backwash cycle: Some historical cases." D. Nunan, V. Berry, R. Berry, eds. *Bringing about change in language education: Proceedings of the international language in education conference*. Hong Kong: Department of Curriculum Studies. The University of Hong Kong, 1994.
- Vernon, P.E. *The measurement of abilities* (2nd ed.). London: University of London Press, 1956.
- Wall, D. Test Impact and Washback. C. Clapham, D. Corson, eds. *Language Testing and Assessment, 7 of The Kluwer Encyclopedia of Language and Education*. Dordrecht: Kluwer Academic, 1997. 291–302.
- . *The impact of high-stakes examinations on classroom teaching. A case study using insights from testing and innovation theory*. Cambridge: University Press. 2005.
- Watanabe, Y. *Washback effect on the Japanese university entrance examinations: classroom-bases research*. PhD thesis, unpublished. Lancaster Univeristy, 1997.
- Weir, C.J. *Communicative language testing*. New York: Prentice-Hall Regent, 1990.
- Wesche, M. "Second language performance testing: The Ontario Test of ESL as an Example." *Language Testing* 4 1 (1987): 28–47.
- Wilczyńska W., Michońska-Stadnik A. *Metodologia badań w glottodydaktyce. Wprowadzenie*. Avalon: Flair, 2010.
- William, D., Black, P. "Meanings and consequences: A basis for distinguishing formative and summative functions of assessment." *British Educational Research Journal* 22 (1996): 537–548.

Wiseman, S., ed. *Examinations and English education*. Manchester: Manchester University Press, 1961.

Wysocka, M. *Profesjonalizm w nauczaniu języków obcych*. Katowice: Wydawnictwo Uniwersytetu Śląskiego, 2003.

Zawadowska-Kittel, E. Efekt zwrotny egzaminu maturalnego z języka angielskiego na proces nauczania i uczenia się (badania pilotażowe). *Języki obce w szkole*, 4. 50–67, 2010.