



Wpływ testu na środowisko edukacyjne

Nick Saville

Dzięki współpracy redakcji JOWS z Cambridge English Language Assessment mamy przyjemność udostępnić Państwu na łamach naszego czasopisma artykuł Nicka Saville'a – jednego z czołowych badaczy, zajmujących się problematyką testowania biegłości językowej. Tekst opisuje model umożliwiający badanie, w jaki sposób ocenianie umiejętności językowych kształtuje konteksty edukacyjne. Tekst w języku angielskim jest poprzedzony krótką syntezą w języku polskim, opracowaną przez dr Agnieszkę Dryjańską, redaktor JOWS.

Badania nad *wpływem testu* mają na celu pokazanie skutków związanych ze stosowaniem testowania i egzaminowania w środowisku edukacyjnym i społecznym. Termin angielski *impact*, tłumaczony na język polski jako *wpływ testu*¹, stanowi niejako rozszerzenie znanego w edukacji pojęcia *washback* – *efekt zwrotny*. O ile *washback* oznacza wpływ testowania na proces nauczania/uczenia się obserwowany w skali mikro – ucznia, nauczyciela, klasy – o tyle *impact* staje się pojęciem nadrzędnym oznaczającym zespół rezultatów przeprowadzania testów i egzaminów obserwowalnych na poziomie społecznym. Celem niniejszego artykułu jest przedstawienie najważniejszych cech obecnie opracowywanego modelu *wpływu testu* oraz wyjaśnienie, w jaki sposób może on być wykorzystany w przypadku egzaminów Cambridge English Language Assessment. Działania zmierzające do wdrożenia tego modelu w różnych

częściach świata mają charakter cykliczny: w każdym cyklu wprowadzane są kolejne zmiany i adaptacje, które okazują się niezbędne w świetle uzyskiwanych danych.

W literaturze przedmiotu już od lat 90. XX wieku analizowane są różne aspekty tzw. *efektu zwrotnego*, jednakże rozważania te nie prowadziły do powstania spójnego modelu uwzględniającego szeroki i złożony kontekst edukacyjno-społeczny, w którym dochodzi do działań ewaluacyjnych. Opracowanie takiego właśnie modelu stanowiło cel projektu realizowanego przez grupę badaczy z Cambridge English Language Assessment. Jedną z najważniejszych teorii uczenia się, stanowiącą podstawę dla prac nad opracowaniem modelu *wpływu testu* przez tych badaczy, jest konstruktywizm – teoria w najbardziej adekwatny sposób opisująca konteksty, w których zachodzi testowanie. Wskazuje ona na wagę takich elementów, jak czynniki afektywne i komunikat zwrotny akcentujący mocne strony ucznia. Mają one wpływ na wzrost motywacji ucznia i lepsze efekty uczenia się.

W projekcie Cambridge English Language Assessment uwzględniono prace Bachmana z roku 1996, który jako pierwszy uznał *wpływ (impact)* za jedną z ważnych cech dobrego testu. Doprowadziło to do przyjęcia modelu VRIP, który

¹ Termin *impact* może być tłumaczony także jako *oddziaływanie testu*, jak proponuje *Angielsko-polsko-słoweński glosariusz terminów z zakresu testowania biegłości językowej* pod redakcją W. Martyniuka, konsultacja naukowa prof. B. Niemierko, prof. H. Komorowska, prof. W. Miodunka, dr J. Magiera, Kraków, 2004. Jednakże w obecnej polskiej literaturze przedmiotu częstsze jest użycie terminu *wpływ*. Należy jednak podkreślić, że definicja pojęcia *impact* prezentowana i dyskutowana przez Cambridge English Language Assessment w niniejszym artykule wykracza poza tradycyjne rozumienie tego pojęcia.

obejmuje także trzy inne cechy: trafność (*validity*), rzetelność (*reliability*) oraz praktyczność (*practicality*) (Saville 2003).

W tym samym roku opracowano wstępny model *wpływu testu*. Obejmuje on cztery wskazania:

- **zaplanuj** (cykliczny proces opracowywania testu);
 - **wspieraj** (uczestników tego procesu w jego realizacji);
 - **informuj** (uczestników o przebiegu procesu);
 - **monitoruj i oceniaj** (przebieg procesu poprzez zbieranie i analizę danych istotnych dla dalszego przebiegu procesu).
- Kolejne projekty zmierzające do opracowania modelu *wpływu testu* wykazały konieczność wyjścia poza tradycyjny kontekst oceniania i nauczania/uczenia się, jakim jest klasa lekcyjna, i uwzględnienia szerszego kontekstu edukacyjno-społecznego. *Wpływ celowy* stanowi kluczowy element tego poszerzonego modelu.

Koncepcja wpływu celowego (testu)

Koncepcja ta polega na opracowaniu takiego systemu oceniania, który będzie charakteryzował się możliwością pozytywnego oddziaływania na środowisko edukacyjne. System taki powinien być opracowany z uwzględnieniem wstępnie przewidywanych rezultatów testowania mogących pojawiać się w konkretnych kontekstach. Model ten obejmuje cztery aspekty: cechy testu, konteksty, rezultaty badane w czasie, metody badawcze i role badaczy.

Jedną z podstawowych cech dobrego testu jest jego trafność. Kluczowym elementem umożliwiającym określenie poziomu trafności analizowanego testu jest opracowanie modelu konstruktów, czyli teoretycznego modelu cechy badanej w procesie testowania. W przypadku testów językowych konstrukt jest badana umiejętność językowa. Jednakże opracowanie właściwego teoretycznego modelu konstruktów nie jest wystarczającym elementem umożliwiającym osiągnięcie założonych rezultatów. Istotne jest włączenie w proces opracowywania testu kryterium *wpływu celowego*, które uwzględnia konteksty społeczne i edukacyjne nauczania/uczenia się i wiąże się z koniecznością współpracy wszystkich uczestników procesu opracowywania testu.

W podejściu do egzaminów i testowania prezentowanym przez Cambridge English Language Assessment kontekst jest definiowany jako dynamicznie oddziałujące na siebie różnorodne elementy tworzące złożone środowisko edukacyjne. Środowisko to obejmuje skalę zjawisk makro, dziejących się kontekście krajowym czy regionalnym, i oznacza politykę edukacyjną, podejmowane reformy edukacyjne, zespół norm kulturowych

i związanych z nimi oczekiwań społecznych wobec edukacji, organizację systemu edukacyjnego, zróżnicowanie geograficzne, ekonomiczne i społeczne pomiędzy regionami oraz skalę zjawisk mikro – dziejących się w kontekście lokalnym – w konkretnych szkołach, klasach, dotyczących indywidualnych uczniów i nauczycieli (Figure 1.)². Kluczowym elementem tego podejścia jest przekonanie, że zależności pojawiające się w konkretnych szkołach i klasach, dotyczące konkretnych uczniów i nauczycieli nie mogą być rozpatrywane w oderwaniu od szerszego kontekstu edukacyjnego i społecznego. Najogólniej rzecz ujmując, skala różnorodności badanych zjawisk wzrasta, gdy oddalamy się od kontekstu makro, zbliżając się do kontekstu mikro (Figure 2.)³. Szczegółowe badania tych zależności, wykorzystujące zarówno metody ilościowe, jak i jakościowe, pozwalają określić, które czynniki ułatwiają, a które utrudniają uzyskanie pozytywnych rezultatów testowania. Niezbędne dane dotyczące kontekstu edukacyjno-społecznego testowania są pozyskiwane w wyniku współpracy między „dostawcą” egzaminów działającym na szczeblu międzynarodowym a lokalnymi „odbiorcami” tych egzaminów. Analiza tych danych dostarcza rozwiązań umożliwiających pokonanie różnego rodzaju trudności pojawiających się na poszczególnych etapach procesu opracowywania testu.

Na etapie praktycznego wdrożenia testu, w docelowych kontekstach i w założonym czasie, szczególnie ważna jest realizacja długoterminowego planu walidacyjnego. Polega ona na cyklicznym analizowaniu rzeczywistych rezultatów testowania – zamierzonych i niezamierzonych, badanie zmienności tych rezultatów w czasie, porównanie rzeczywistych rezultatów z przyjętymi założeniami i wprowadzanie odpowiednich zmian, których rezultaty są znów w ten sam sposób analizowane i weryfikowane.

Konkluzja

Najważniejszą cechą opracowanego modelu *wpływu celowego* (testu) jest możliwość wprowadzania zmian w celu poprawy rezultatów i zmniejszenia pojawiających się złych skutków testowania. Samo przewidywanie rezultatów i porównywanie ich z rezultatami rzeczywistymi nie jest wystarczające, jeżeli nie implikuje zmian prowadzących do poprawy rzeczywistych skutków. Dlatego też ważnym elementem procesu opracowania testu jest zaplanowanie etapu „zarządzania zmianą”, który umożliwia wprowadzanie konkretnych modyfikacji do procesu testowania.

² Diagram (Fig. 1.) dostępny w tekście oryginalnym artykułu.

³ Ibidem.



Applying a model for investigating the impact of language assessment within educational contexts: the Cambridge English approach

In Research Notes 42 (2010), I explained why Cambridge English Language Assessment as an international test provider needs a model to guide its work in investigating the impact of its examinations. In this article I set out some features of the model now being developed and explain how it can be applied in the case of the Cambridge English examinations. The operational practices needed to implement this approach are being introduced incrementally and are being adapted and revised in light of experiences in conducting projects which are now underway in many parts of the world.

Impact research within Cambridge English

Impact research investigates and seeks to understand the effects and consequences which result from the use of tests and examinations in educational contexts and throughout society. As a field of enquiry it appeared in the language testing literature as an extension of washback in the 1990s. (See Cheng, Watanabe and Curtis 2004 for an overview of washback.) The PhD theses of Wall (2005), Cheng (1997, 2005) and Green (2007) published in the Studies in Language Testing series, looked at different aspects of washback and extended the earlier work of Hughes (1989) and Bailey (1996). While these studies inevitably touched on considerations related to impact, none proposed a comprehensive model which would allow complex relationships to be examined across wider educational and societal contexts. This has been the aim of the team working in Cambridge English.

The origin of the Cambridge English approach dates back to the early 1990s and to the time when the current test development and validation strategies were first introduced. In those early stages, Bachman's work was influential as he was the first to present impact as a 'quality' of a test which should be integrated within the overarching concept of test usefulness (Bachman and Palmer 1996). Following his lead, Cambridge English also introduced impact as one of the five essential qualities, which together with validity, reliability, practicality and quality comprise the VRIPQ features of a test.

By conceptualising impact within VRIPQ-based validation processes from the start, there was an explicit attempt to integrate impact research into routine procedures for accumulating validity evidence. Subsequent work on impact has been framed by these considerations and since the initial stage it has been recognised that a proactive approach is needed to achieve intended effects and consequences.

In 1996, Milanovic and Saville proposed an early model of test impact which was explicitly designed to meet the needs of Cambridge English. They proposed four maxims as follows:

- Maxim 1 PLAN
Use a rational and explicit approach to test development;
- Maxim 2 SUPPORT
Support stakeholders in the testing process;
- Maxim 3 COMMUNICATE
Provide comprehensive, useful and transparent information;
- Maxim 4 MONITOR and EVALUATE
Collect all relevant data and analyse as required.

These maxims were designed to capture key principles and to provide a basis for practical decision-making and action planning – and they still remain central to the Cambridge English approach today (see Section 4.4 in Cambridge English Principles of Good Practice: Quality Management and Validation in Language Assessment (2013).

Under Maxim 1, Cambridge English endeavours to develop systems and processes to plan effectively using a rational and explicit model for managing the test development processes

in a cyclical and iterative way. It requires regular reviews and revisions to take place and for improvements to be made when necessary (Cambridge English 2013:18–22, Saville 2003:57–120).

Maxim 2 focuses on the requirement to support all the stakeholders involved in the processes associated with international examinations. This is an important aspect of the approach because examination systems only function effectively if all stakeholders collaborate to achieve the intended outcomes.

Maxim 3 focuses on the importance of developing appropriate communication systems and of providing essential information to the stakeholders (Cambridge English 2013:12–14).

Maxim 4 focuses on the essential research requirement to collect as much relevant data as possible and to carry out routine analyses as part of the iterative model (noted under Maxim 1). The nature of the data needed to investigate impact effectively and how it can be collected, analysed and interpreted under operational conditions has become an increasingly important part of the model in recent years.

Three major impact studies were also carried out between 1995 and 2004. Project 1 was the survey of the impact of IELTS (International English Language Testing System). This project helped conceptualise impact research including the design and validation of suitable instruments. Project 2 was the Italian Progetto Lingue 2000 impact study and was an application of the approach within a single macro educational context. These two projects are described in detail by Hawkey (2006). Project 3 was the Florence Language Learning Gains Project (FLLGP). Still within Italy, this project was an extension and re-application of the model within a single school context (i.e. at the micro level). It focused on individual stakeholders in one language teaching institution, namely teachers and learners preparing for a range of English language examinations at a prestigious language school in Florence. The complex relationships between assessment and learning/teaching in a number of language classrooms, including the influence of the Cambridge English examinations, were examined against the wider educational and societal milieu in Italy. The micro level of detail, as well as the longitudinal nature of the project conducted over an academic year, were particularly relevant in this case (Saville 2009).

Based on an analysis of these projects, I have proposed a meta-framework designed to provide a more effective model for conducting impact research under operational conditions (Saville 2009). I suggest that by implementing this framework more systematically, ‘anticipated impacts’ can be achieved more effectively and well-motivated improvements to the examination systems can be identified and put into place. Aspects of this approach are represented in the impact studies reported in this issue and are focused on in the second part of this paper under the concept of impact by design.

The concept of impact by design

Impact by design is a key feature of the expanded impact model. It starts from the premise that assessment systems should be designed from the outset with the potential to achieve positive impacts and takes an *ex ante* approach to anticipating the possible consequences of using the test in particular contexts.

In the final part of this paper, the following four points which are central to the model are discussed:

- test features (constructs and delivery systems);
- contexts;
- outcomes over time – the timeline;
- research methods and roles of researchers.

Test features (constructs and delivery systems)

Impact by design builds on Messick’s (1996) idea of achieving ‘validity by design as a basis for washback’. The importance of the rational model of test development and validation with its iterative cycles is a necessary condition for creating construct valid tests and for the development of successful systems to support them (cf. Maxim 1). Adequate specification and communication of the focal constructs is crucial for ensuring that the test is appropriate for its purpose and contexts of use and to counter threats to validity: construct underrepresentation and construct irrelevant variance (Messick 1996:252).

Insights from socio-cognitive theory underpin contemporary theories of communicative language ability, language acquisition and assessment (cf. the socio-cognitive model (Cambridge English 2013:25–27, Weir 2005)) and are also helpful in understanding how language learning and preparation for examinations takes place in formalised learning contexts, such as classrooms.

While appropriate construct representation is a necessary condition for achieving the anticipated outcomes, it is not sufficient and impact by design highlights the importance of designing and implementing assessment systems which explicitly incorporate considerations related to the social and educational contexts of learning/teaching and test use. This relates to the need for effective communication and collaboration with stakeholders, as noted in the original Maxims 2 and 3 and incorporated into the Principles of Good Practice, Section 2 (Cambridge English 2013).

Contexts

Understanding the nature of context within educational systems and the roles of stakeholders in those contexts are clearly important considerations for Cambridge English – see Saville (2003:60) for a discussion of stakeholders.

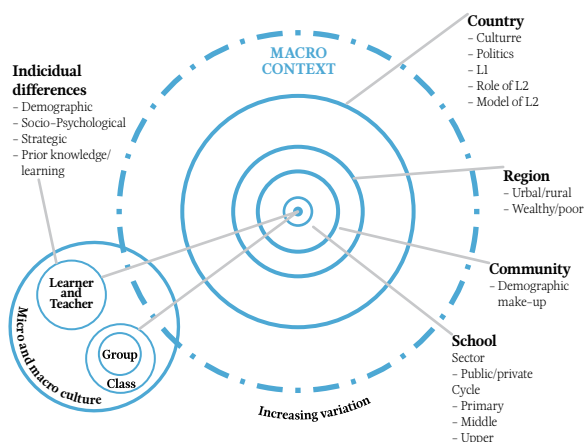


Figure 1: Context in education = a complex dynamic system

It is now widely recognised that educational processes (see Figure 1) take place within complex systems with dynamical interplay between many sub-systems and ‘cultures’ and so an understanding of the roles of stakeholders as participants is a critical factor in bringing about intended changes (e.g. Fullan 1993, 1999, Thelen and Smith 1994, Van Geert 2007).

In conducting impact research the aim is to understand better the interplay between the macro and micro contexts within the society where the tests are being used and to determine which elements facilitate or hinder the desired outcomes. In general, diversity and variation increases as one moves from the general milieu within a country or region (the macro context) to specific schools and ultimately to the

individual participants within classrooms (the multiple micro contexts at the local level involving schools, classes/groups and individual teachers and learners).

Figure 2 diagrammatically shows a school context embedded in a wider milieu with a teacher interacting with groups of learners in a particular classroom. The external influences include the general features of the milieu, as well as specific educational factors such the curriculum and syllabus and the need to produce examination results which are used outside of the school context.

It is therefore important to develop methods to understand both the general context as well as specific local cases, including dynamics which affect learning in classrooms. This points to the need to use both quantitative and qualitative data collection methods (see below).

In understanding the macro contexts into which international examinations are introduced (e.g. as part of educational reforms or innovations), it is important to focus on key factors related to the following:

- the political regime and its approach to educational reforms;
- the role of educational reforms within wider socioeconomic policies;
- cultural norms and expectations in relation to education generally, and attitudes towards language education (and towards English specifically in the case of Cambridge English);
- the educational system and how it is organised (e.g. compulsory education and the nature of the educational cycles; private vs. public schools; role of standardised assessment, etc.);
- broad differences between geographical regions and socioeconomic groups.

Collaboration between an international examination provider and local users is essential in order to capture relevant data and to shed light on such contextual parameters. Many dilemmas which arise in assessment contexts can only be dealt with if a wide range of local stakeholders agree to manage them in ways which they jointly find acceptable; the challenge is to get the relevant stakeholders working together effectively to agree what needs to be done to achieve the intended outcomes.

Outcomes over time – the timeline

It is essential to know what happens when a test is introduced into its intended contexts of use and this should constitute

a long-term validation plan (cf. Maxims 1 and 4). Anticipating and managing change over time within specific contexts is therefore central to this concept and it means that appropriate consideration of timescales and the timeline for implementation (often involving several phases) are central to the design of impact studies. In impact research designs there is nearly always a fundamental need to collect comparative data, and therefore to develop research designs which can be carried out in several phases over an extended period of time or replicated in several different contexts.

Similarly, effects and consequences – intended and unintended – usually emerge over time given that contexts of use are not uniform and are subject to change, e.g. as a result of localised socio-political and other factors. Impact by design is therefore not strictly about prediction; a more appropriate term might be ‘anticipation’. In working with stakeholders, possible impacts on both micro and macro levels can be anticipated as part of the design and development process, and where potentially negative consequences are anticipated, remedial actions or mitigations can be planned well in advance.

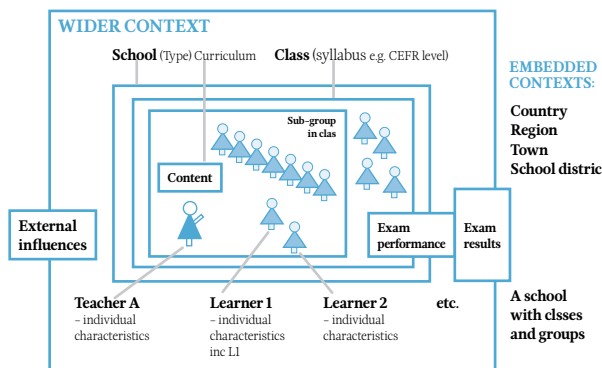


Figure 2: Multiple classroom context where teaching and learning take place

Research methods and roles of researchers

Contemporary theories of knowledge and learning have played a prominent role in developing Cambridge English model of impact and the search for a ‘paradigm worldview’ (epistemology and ontology) which provides an effective conceptualisation and has drawn on relevant theories in the social sciences. A ‘realist’ stance now underpins Cambridge English approach, drawing on ‘critical realism’ (e.g. Sayer 1984, 2000) and contemporary views on pragmatism.

Constructivism is also important for the re-conceptualisation of impact for two reasons: first because contemporary approaches to teaching and learning in formal contexts now appeal to constructivist theories; secondly, because it is most appropriate to finding out ‘what goes on’ in contexts of test use. From the learner’s perspective, affective factors are vital for motivation and feedback that highlights strengths positively tends to lead to better learning (i.e. learning oriented assessment). These considerations are relevant in designing language assessment systems which have learning oriented objectives and a concern in impact research is whether these objectives have been met effectively.

The current model of impact looks to ‘real world’ research paradigms to provide tools which can shed light on what happens in testing contexts, including mixed methods and quasi-experimental designs. Case studies are especially useful for investigating impact at the micro level and for understanding the complexities of interaction between macro level policies and implementation in local settings. Without such methods it is difficult to find out about and understand how the interaction of differing beliefs and attitudes can lead to consensus or to divergence and diversity.

Mixed method research designs are becoming increasingly relevant to addressing impact research questions. Creswell and Plano Clark (2011:69) discuss six prototypical versions of mixed method research designs which seek to integrate qualitative and quantitative data in parallel and sequential ways and these are becoming central to the Cambridge English approach, as illustrated by the studies reported in this issue.

The Cambridge English ‘impact toolkit’ of methods and approaches is now being used to carry out analyses of both large-scale aggregated data, as well as micro analyses of views, attitudes and behaviours in local settings (as in the earlier case of the Progetto Lingue 2000 impact study reported by Hawkey 2006). Quantitative analysis of macro level group data allows us to capture overall patterns and trends, while the qualitative analysis of multiple single cases enables the research team to monitor variability in local settings and to work with the ‘ecological’ features of context. It is the integration of both analyses to provide the insights and interpretations which is particularly important.

Finally it is important to highlight the make-up of the impact research teams; where possible, the team should

comprise both Cambridge-based staff with appropriate skills in research design and analysis, as well as local researchers who may be 'participants' in the teaching/learning context itself and who bring a deeper understanding of the educational and cultural context which is under investigation. Again this is illustrated in the studies reported in this issue, including Gu and Saville working jointly with other participants in the Chinese context.

Conclusion

The ability to change in order to improve educational outcomes or mitigate negative consequences associated with the examinations is ultimately the most important dimension of the impact by design model. Anticipating impacts and finding out what happens in practice are not enough if improvements do not occur as a result. Being prepared to manage change is therefore critical to a theory of action. In working closely with the stakeholders in their own contexts, this approach is now providing us with the necessary tools to determine what needs to be done and when/how to do it.

References

- Bachman, L., Palmer, A. (1996) *Language Testing in Practice*. Cambridge: Cambridge University Press.
- Bailey, K.M. (1996) Working for Washback: A Review of the Washback Concept in Language Testing. W: *Language Testing*, nr 13 (3), 257-279.
- Cambridge ESOL (2011) *Principles of Good Practice: Quality Management and Validation in Language Assessment* [online] [dostęp 8.04.2015] <<http://www.cambridgeenglish.org/images/22695-principles-of-good-practice.pdf>>.
- Cheng, L. (1997) *The Washback Effect of Public Examination Change on Classroom Teaching: An Impact Study of the 1996 Hong Kong Certificate of Education in English on the Classroom Teaching of English in Hong Kong Secondary Schools* [niepublikowana praca doktorska]. University of Hong Kong [niepublikowana praca doktorska].
- Cheng, L. (2005) Changing Language Teaching through Language Testing: A Washback Study. W: *Studies in Language Testing*, nr 21. Cambridge: UCLES/Cambridge University Press.
- Cheng, L., Watanabe, Y., Curtis, A. (red.) (2004) *Washback in Language Testing: Research Contexts and Methods, Mahwah*. Nowy Jork: Lawrence Erlbaum Associates.
- Creswell, J.W., Plano Clark, V.L. (2011) *Designing and Conducting Mixed Methods Research*. Thousand Oaks. California: Sage.
- Fullan, M. (1991) *The New Meaning of Educational Change*. London: Cassell.
- Fullan, M. (1993) *Change Forces: Probing the Depths of Educational Reform*. London: The Falmer Press.
- Fullan, M. (1999) *Change Forces: The Sequel*. London: The Falmer Press.
- Green, A. (2003) *Test Impact and EAP: A Comparative Study in Backwash Between IELTS Preparation and University Pre-session Courses* [niepublikowana praca doktorska]. University of Surrey.
- Green, A. (2007) IELTS Washback in Context: Preparation for Academic Writing in Higher Education. W: *Studies in Language Testing*, nr 25. Cambridge: UCLES/Cambridge University Press.
- Hawkey, R. (2006) Impact Theory and Practice: Studies of the IELTS Test and Progetto Lingue 2000. W: *Studies in Language Testing*, nr 24. Cambridge: UCLES/Cambridge University Press.
- Hughes, A. (1989) *Testing for Language Teachers*. Cambridge: Cambridge University Press.
- Messick, S. (1996) Validity and washback in language testing. W: *Language Testing*, nr 13 (3), 241-256.
- Milanovic, M., Saville, N. (1996) *Considering the Impact of Cambridge EFL Examinations* [raport wewnętrzny]. Cambridge: Cambridge ESOL.
- Saville, N. (2003) The Process of Test Development and Revision within UCLES EFL. W: C. Weir, M. Milanovic (red.) *Continuity and Innovation: Revising the Cambridge Proficiency in English Examination 1913-2002*. *Studies in Language Testing*, nr 15. Cambridge: UCLES/Cambridge University Press, 57-120.
- Saville, N. (2009) *Developing a Model for Investigating the Impact of Language Assessment within Educational Contexts by a Public Examination Provider* [niepublikowana praca doktorska]. University of Bedfordshire
- Saville, N. (2010) Developing a Model for Investigating the Impact of Language Assessment. W: *Research Notes*, nr 42, 2-8.
- Sayer, A. (1984) *Method in Social Science: A Realist Approach*. London: Routledge.
- Sayer, A. (2000) *Realism and Social Science*. London: Sage.
- Thelen, E., Smith, L.B. (1994) *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge, Massachusetts: The MIT Press.
- Van Geert, P. (2007) Dynamic Systems in Second Language Learning: Some General Methodological Reflections. W: *Bilingualism: Language and Cognition*, nr 10, 47-49.
- Wall, D. (1999) *The Impact of High-stakes Examinations on Classroom Teaching: A Case Study Using Insights From Testing and Innovation Theory* [niepublikowana praca doktorska]. Lancaster University.
- Wall, D. (2005) The Impact of High-Stakes Examinations on Classroom Teaching: A Case Study Using Insights from Testing and Innovation Theory. W: *Studies in Language Testing*, nr 22. Cambridge: UCLES/Cambridge University Press.
- Weir, C.J. (2005) *Language Testing and Validation: An Evidence-based Approach*. Basingstoke: Palgrave Macmillan.

* Artykuł przedrukowany z *Research Notes Issue nr 50*.

dr Nick Saville

Członek kadry kierowniczej w Cambridge English, odpowiedzialny za koordynację prac grupy badawczej ds. walidacji. Bierze czynny udział w inicjatywach europejskich, m.in. w rozwoju *Europejskiego systemu opisu kształcenia językowego* (ESOKJ) oraz reprezentuje Cambridge English w ALTE (Association of Language Testers in Europe). Członek zespołu Cambridge English odpowiedzialny za koordynację programu *English Profile* oraz współtwórca Instytutu Automatyzacji Nauczania i Testowania na Uniwersytecie Cambridge. Były redaktor naczelny periodyku *Language Assessment Quarterly* (od 2004 r.) oraz współredaktor serii *English Profile Studies*.