

Dorota Bednarek  <https://orcid.org/0000-0003-4777-2588>

SWPS University of Social Sciences and Humanities

e-mail: [dbednarek@swps.edu.pl](mailto:dbednarek@swps.edu.pl)

Katarzyna Cybulska-Gómez de Celis  <https://orcid.org/0000-0002-8883-3175>

Siedlce University of Natural Sciences and Humanities

e-mail: [katarzyna.cybulska-gomez@uph.edu.pl](mailto:katarzyna.cybulska-gomez@uph.edu.pl)

## Towards research-literate language educators: Unveiling difficulties and delineating pathways of evidence-informed practice

### Abstract

The aim of our text is to show the potential benefits of developing research literacy of teachers in general and language teachers in particular. We argue that research-literate teachers may effectively draw upon empirical studies and make informed decisions in the classroom, which has the potential to optimise students' benefits. We emphasise the fact that, by referring to their experiential knowledge, teachers can also assume the role of active researchers who communicate their observations and validate particular teaching techniques or activities for different contexts and needs. By attempting to clarify basic misunderstandings, we see the need for 'evidence-informed' practice in teaching, which stems from an 'evidence-based' approach popular in medicine, psychology and psychotherapy. Finally, we demonstrate the suitability of various designs of empirical studies to answer different types of research questions, which we exemplify by high-quality studies from language education. Concluding, we call for tighter cooperation and bidirectional exchange of the know-how between teachers and academics.

**Keywords:** research literacy, evidence-informed practice, evidence-based practice, efficacy research in education, language teachers, language education

## Introduction

The idea of teachers' research literacy (e.g. Neufeld, 1990), followed by proposals of research literacy training programs (e.g. Tuñón, 2002; Amir, Mandler, Hauptman, Gorev, 2017) has a few decades of history, and yet, we argue that informing teaching practice through scientific findings in many cases seems to be still just an aspiration for the policymakers rather than everyday teaching practice (see Center on Education Policy, 2019).

Our main objective is to support the development of research literacy among teachers and educators. We do so by offering practical guidelines on the evaluation of different empirical research methods. The understanding of the advantages and disadvantages of particular research designs may help teachers critically select valuable empirical evidence to inform their classroom decisions.

Data shows that, in many cases, teachers tend to obtain information about classroom practices from other teachers rather than from research (Cooper, Klinger, McAdie, 2017) and express minor interest in becoming acquainted with the empirical evaluation of teaching methods (Sari, 2006; Mausethagen, Raaen, 2017). Due to the benefits of evidence-informed practice in various applied disciplines and the availability of multiple empirical studies on foreign language teaching, there are potential gains resulting from consulting scientific data by language teachers. However, an insufficient focus on research literacy in teacher education curricula or further professional development as well as certain misrepresentations of the empirical paradigm in academic discussions in the area of pedagogy may add to the relatively limited teachers' interest in consulting and conducting research in their everyday practice.

Additionally, we shed light on the complexity and challenges of promoting research literacy among language teachers by first exploring the medical origins of the so-called 'evidence-based' concept of practice and its critique. We explain the difference between the type of practice based on assumptions versus the one that is informed through the empirical findings. Next, we address the main objections expressed in pedagogical theoretical discussions towards the implementation of approaches which extensively draw upon empirical research and measurement in education. In this part, we analyse such issues as expected reduction of a practitioner's autonomy, simplified vision of a patient/client/student, the problem of standardised testing and randomised controlled trials. After explaining basic concepts and attempting to refute some mistargeted criticism, we emphasise the need for greater research literacy among teachers and policymakers via showing the importance of empirical verification of teaching methods or techniques on the one hand and indicating some unsuccessful applications of empirical studies to educational policies on the other.

In the main part of this text, we present a wide spectrum of different research methods, showing their relative usefulness for different aims, to motivate teachers not only to critically evaluate available sources of data but also to take an initiative by promulgating their own observations reflecting their experiential knowledge and undertaking their own empirical studies. These guidelines are based on the American Psychological Association materials, but they are also useful in the language teaching context.

In conclusion, we call for training of teachers in research literacy and establishing collaborative links between teachers and research centres. In our opinion, an evidence-informed approach and the monitoring of learning outcomes in varied contexts may benefit not only students but also teachers while rewarding them with a sense of agency.

## **A common platform between research literacy and evidence-informed practice**

Research literacy may be defined as one's ability 'to engage with research in order to assess its utility and ripeness for adaptation to context. It is not about an unthinking acceptance of received opinion. It involves critical scrutiny of evidence' (Waring, Evans, 2015: 18). In other words, it refers to the ability to understand and critically evaluate empirical evidence as well as being able to discern the strong and weak points of published studies. It involves a readiness to conduct one's empirical endeavours and share one's own observations with others. We believe that well-prepared language teacher educators should play a leading role in promoting research literacy through teacher training programs for pre-service and in-service language teachers. By acting as facilitators, teacher educators shall provide practitioners with ways to access empirical publications and guide them on how to select and implement insights from valuable academic research.

Research-literacy strongly relates to the idea of practitioners drawing not only on authoritative sources and their own professional experience but also on empirical evidence. To give an example from psychotherapy, '[i]deally, practitioners who actively employ EBPs [evidence-based practices] save time, money, and resources by avoiding treatments with little or questionable effectiveness for their patients' (Cook, Schwartz, Kaslow, 2017: 538). Nevertheless, the introduction of such a strong empirical foundation into disciplines that emphasise the individual human experience as the main source of professional decisions inevitably leads to tensions.

Nowadays, policymakers in education vigorously promote the idea of the so-called 'research-informed', 'evidence-informed' or 'evidence-based' practice in teaching (e.g. ResearchEd, n.d.; Education Endowment Foundation, 2020). While some authors focus on the differences between these terms (for discussion see: Nelson, Campbell 2017), others diminish the terminological importance (e.g. Woodbury, Kuhnke, 2014). The prominence of the experiential knowledge in integrating and selecting insights from empirical data (cf. Chalmers, 2005; Nevo, Slonim-Nevo, 2011; Sharples, 2013) seems to make the 'evidence-informed' term more inclusive and up-to-date rather than 'evidence-based' although we may not shun from the latter due to its historical presence, significance and acceptance in the contexts of medicine, psychology and psychotherapy. Above all, however, '[t]he terminology is less important than the approach' (Woodbury, Kuhnke, 2014: 21) and our main aim is to highlight the distinction between two types of teaching practice: one supported and informed by empirical evidence and focused on monitoring teaching effects in a given context versus the second type of practice which neglects these two aspects. Just like 'evidence-based psychotherapies are associated with higher quality

and more accountability’ (Cook, Schwartz, Kaslow 2017: 539), we believe that the same may be true for evidence-informed teaching practice.

## A challenging application of a ‘medical model’ to education

One of the excellent explorations of possible applications of an evidence-informed approach to education may be exemplified by Becher and Lefstein’s (2020) reflections in *Teaching as a Clinical Profession: Adapting the Medical Model*. The paper examines the challenges stemming from an adaptation of this paradigm to the social context characterised by the interactional character of teaching, showing that it is neither obvious nor easy. The authors discuss analogies and discrepancies between medical and teaching practice and indicate that three issues have to be addressed: 1) parallels between medical diagnosis and needs analysis of learners; 2) the applicability of the concept of treatment (medical intervention) to the context of teaching process (teaching strategy); and 3) changeability of teachers’ actions resulting from the ongoing monitoring and constant adjustment to the needs and obtained results. Their observations precisely pinpoint the basic obstacles to evidence-informed practice in teaching. They look for ways of successful implementation of the ‘medical model’ in education to promote the selection of practice of proven efficacy over the unverified one, an approach which could and should be more widely adopted by teachers. In our opinion, especially in foreign language teaching, there is a growing number of good-quality studies comparing the effects of different teaching techniques or classroom activities, suggesting that language teachers can and should be in the vanguard of teaching informed by evidence.

Modern medicine, just like psychotherapy and education, evolved from practice based on the premises representing early non-scientific understanding of physiology, mind or learning processes. This type of attitude may be illustrated by a Latin sentence: *Melius anceps remedium quam nullum*, which means that it is better to try a dubious remedy rather than to do nothing. Long before any reliable knowledge about biological and mental functions was available, people had exercised medicine, mental help, and teaching. This explains the persistence of harmful treatments in medicine, such as bloodletting (Burch, 2013), damaging psychotherapies, such as reparatory therapy for non-heterosexuality (Van Zyl, Nel, Govender, 2017), or counterproductive educational practices, including the use of corporal punishment and intimidation (Civil Rights Project, 2019). Generations of practitioners were introduced to their profession by mentors who attributed health problems to unsubstantiated causes (see: Lagay, 2002 – somatic illness linked to the disequilibrium of the vapours; Ng, 1999 – mental issues attributed to the wandering womb).

The erratic theories, lacking empirical verification, were followed by dangerous treatments devoid of monitoring of effects. Let us provide a few very expressive examples. George Washington, among many others, died after having 40% of his blood spilt due to the practice resulting from an unscientific medical theory, the so-called ‘humoral physiology’ (Burch, 2013). The unproven assumption that autism was caused by the so-called ‘refrigerator mother’ (mother’s unemotional attitude towards her child, Kanner, 1943) led

to a policy (which lasted until the 1980s) of compulsory separation of autistic children from their families causing not an improvement but the aggravation of the children's condition (Sousa, 2011). After empirical studies provided verified objective data to explain the nature of the patients' suffering, there was still a need for another revolution, and this time it was not in the conceptual but in the practical sphere. Even though the understanding of physiological or psychological processes promotes practice that adequately addresses the needs, each intervention requires to be carefully examined using empirical research to guarantee that the chosen remedy renders expected results and has no grave side effects. Although it was not a sudden discovery but rather a long-lasting process, the idea of 'evidence-based medicine' was formulated in the 1980s (see: Nordenstrom, 2007) as a responsible practice of applying remedies with empirically proven effects and controlling the outcomes for the patient. Practice informed by evidence means that each treatment procedure needs to obtain reliable verification of beneficence and non-maleficence for a given category of patients with a given type of medical condition. In consequence, patients receive treatments that, due to the best current knowledge supported by evidence, are most likely to help them without provoking serious side effects.

## Dealing with misconceptions and controversies

The promotion of practice rooted in empirical data within traditionally humanistic disciplines may lead to controversies, some of which are justified, while others are misplaced.

### Does empirical verification reduce the role of a practitioner?

The main misconception regarding the practice that prioritises empirical evidence may consist in believing that it reduces a practitioner's job to blindly applying given treatments or teaching strategies as if they were taken from a 'cookbook' (see: McKnight, Morgan, 2019). Similar fear was popular in the field of psychotherapy; '[t]here is a misperception that evidence-based psychotherapies are mere "cookbook" practice instructions that force clinical professionals to replace their judgment with "manualized" procedures' (Cook, Schwartz, Kaslow, 2017: 539). As we are arguing below, such opinions misrepresent the idea of approaches that draw upon empirical evidence; however, they may portray the risks of their *incompetent implementation*.

The model of practice accepted in modern medicine, psychology and psychotherapy is based on rigorous empirical research and involves at least three separate stages: 1) objective assessment of the patient's condition, 2) the choice and the application of the best treatment (the practitioner's decision supported by known efficacy and risk), and 3) outcome monitoring in each individual case and eventual treatment modification by a conscious practitioner. We believe that this model could be successfully adapted to education. The three stages would be represented by 1) students' needs evaluation, 2) decisions regarding teaching method or technique / classroom management strategy / other teaching activity, and 3) outcome monitoring followed by the modification of the practitioner's earlier decision if necessary. Only in a narrower sense, the term 'evidence-based' medicine,

psychology or psychotherapy is reserved for the middle phase, i.e. treatment assignment. Such a narrow understanding of practice based on empirical evidence led to criticism that it may result in the mechanisation of a complex decision-making process and ‘suppress clinical freedom’ (see: Sackett et al., 1996: 71 for reference). However, the core idea of evidence-based medicine has always been, as it is in the case of evidence-informed practice in education, to promote individual decisions based on the best available data. As stated over 20 years ago:

Evidence based medicine is the conscientious, explicit and judicious use of **current best evidence** in making decisions about the care of individual patients. The practice of evidence based medicine means **integrating individual clinical expertise with the best available external clinical evidence** from systematic research (Sackett et al., 1996: 71; emphasis given by the authors of this text).

In his discussion of the current use of evidence in social policy and practice, Sharples also argues that ‘[e]vidence-based practice is [...] about **integrating professional expertise** with the best external evidence from research to improve the quality of practice’ (Sharples, 2013: 7; emphasis given by the authors of this text). Actually, while conditioning professional decisions upon empirical evidence does not reduce the agency of an expert, it also promotes the individualisation of the decision-making process and reinforces the role of professional expertise,

By individual clinical expertise we mean the proficiency and judgement that individual clinicians acquire through clinical experience and clinical practice. Increased expertise is reflected in many ways, but especially in more effective and efficient diagnosis and in more thoughtful identification and compassionate use of individual patients’ predicaments, rights and preferences in making clinical decisions about their care. [...] Good doctors use both individual clinical expertise and the best available external evidence, **and neither alone is enough** (Sackett et al., 1996: 71–72; emphasis given by the authors of this text).

## **Does empirical verification lead to a reductionist view of a student?**

Another misconception of the approaches that prioritise outcome monitoring and informing professional decisions through empirical evidence is that they impose the simplified and reductionist vision of a patient, client, or student. Although there may be some malpractice, the approaches sustaining the link between practice and evidence, as it may be seen in the examples of medicine, psychology or psychotherapy, embrace the individual, not just as a biological organism but as a whole unique person. Data shows that a patient’s values and the resulting preference for treatment are significant factors in general treatment outcomes, not just in terms of organ improvement but also for the general patient’s well-being (for meta-analysis see e.g. Lindhiem et. al., 2014; Preference Collaborative Review Group, 2009). Therefore, medicine based on empirically verifiable data requires a holistic approach to the patient, and contrary to the sceptics’ views, it neither suppresses the clinical freedom of a practitioner nor reduces individuals to their organ function. Actually, the reverse is true, and this approach has revolutionary consequences for professional autonomy: phy-

sicians cannot be merely technicians applying policies or medical procedures; they must view a patient as a person who reacts individually at the physiological and mental level. This attitude is even better expressed by the advocates of the term ‘evidence-informed’ practice who offer a wider perspective emphasising that ‘[t]he client, not the best evidence, is at the centre’ (Nevom Slonim Nevo, 2011: 15) and ‘the evidence will be considered along with a host of other considerations taken **in equilibrium** by an experienced and imaginative practitioner’ (ibidem, 2011: 10; emphasis given by the authors of this text).

### **Does empirical verification result in inflexible practice?**

Medicine, psychology and psychotherapy based on evidence oblige practitioners not only to use their professional expertise but also to monitor the outcomes for every patient and adjust the treatment in the light of the lack of expected results. Becher and Lefstein (2020), discussing the similarities between the ‘medical model’ and teaching, indicate the centrality of the client (student/patient), the necessity of individual needs assessment (learning needs evaluation/health examination), and the inclusion of modulating factors (social context, individual differences, special needs/comorbid conditions, patient values) in these both disciplines. In their analysis, the most striking difference between education and medicine consists in ‘disciplined improvisation’ inevitably practised by teachers applying their teaching strategies in a constantly changing social context of a classroom. Indeed, teachers, probably much more than doctors, have to constantly adapt their actions to the feedback from their students. Here, as Nevo and Slonim Nevo (2011: 16) argue with respect to evidence-informed practice, ‘use of the evidence in treatment must be attuned to the client and therefore flexible and dynamic’. In other disciplines based on empirical data, like psychotherapy, and group psychotherapy specifically, practitioners have to introduce ‘disciplined improvisation’ with at least the same intensity as teachers, and they include social and emotional context into their decision-making. ‘Applying evidence-based principles ensures that providers use the best existing evidence as a starting framework, while simultaneously affording them flexibility to individualize treatment’ (Cook, Schwartz, Kaslow, 2017: 538). Flexibility is thus the core of practice that is either based on or informed by research in any discipline.

### **Does empirical verification promote standardised testing in education?**

Yet another mistrust of the extensive use of empirical research specifically in pedagogy and language pedagogy relates to the problem of standardised testing. Undoubtedly, there are some rightful concerns about how to operationalise the complexity of teaching, educational aims and outcomes, and which measurable variables should be applied; in other words, the question is how to create good quality evidence in such a complex context as education. One of the worries results from confounding practice drawing upon empirical research, with controversial standardised testing in education and language education. We assume that policymakers turned to standardised testing because tests are relatively cheap and easy to administer and score. Standardised tests of cohorts of pupils,

popular in education, bring actually low-quality evidence with vague research questions and inconclusive results (Wiliam, 2010), which in isolation should not serve as a base for evidence-informed practice (see the next part of our text). Not only are standardised tests in education unrelated to the idea of evidence-based or evidence-informed practice, but they are also examples of measurement not fulfilling high demands for outstanding studies. In order for practice to be well informed by evidence, it requires not massive but good quality data.

The criticism in pedagogy over negative side-effects of extensive testing is well-grounded in evidence. For example, Harlen and Deakin-Crick (2002) performed an excellent systematic review of the impact of tests and other standardised forms of assessment on students' motivation for learning, analysing 19 good quality, reliable studies. The results show the detrimental impact of testing in several areas crucial for the quality of education and students' personal development. As tests are ideal to assess repetitive tasks, they encourage teachers to change practice, neglecting higher-order skills (necessary for open-ended problems, including critical thinking), and to emphasise lower-order skills in order to boost 'good' test results. A worrisome effect was observed through a shift in the quality of motivation for learning: students intensively tested lose mastery-oriented intrinsic motivation (proven to be associated with better learning results) for extrinsic motivation to perform well in a test. Finally, to name just a few findings, it was shown that students with a history of low-test results develop lowered self-image and increased anxiety. The detrimental effect of learning for the sake of passing tests is described in the literature on teaching English as a Foreign Language [EFL]. For example, Barnes (2016) analysed the negative 'washback effect' (see: Alderson, Wall, 1993; Bailey, 1996), consisting in students getting good marks on tests and failing to acquire the necessary language skills due to their teachers' reliance on TOEFL iBT textbooks and their concentration on developing language skills tested by TOEFL iBT.

The profound analysis of these phenomena is beyond the scope of this article; however, it is important to bear in mind that extensive standardised testing leads to several detrimental effects. The opposition in pedagogy against education policies appropriated by simplified measurements is justifiable and based on empirical evidence. Policies inspired by research are not in line with standardisation; to the contrary, if the policies were indeed inspired by evidence, standardised tests would be restrained. In particular, the so-called testing for accountability, which links average school results in standardised tests with financial consequences, does not meet the requirements of policy drawing upon empirical data (see Wiliam, 2010).

## **Does empirical verification reduce scholarship to randomised controlled trials?**

Another source of scepticism in pedagogy regarding the approach accepted in medicine, psychology and psychotherapy may be attributed to misconceptions about the imposition of research methodology. Those who support evidence-based practice are accused of discrediting any other form of scholarship if not based on randomised controlled trials, and randomised controlled trials are seen as inapplicable to serve pedagogical contexts.



Unfortunately, labelled as the ‘gold standard’, randomised controlled trials attract unjustified criticism, while they are just one of many different valuable sources of data to be drawn upon as we will show later in this text. Development of evidence-informed pedagogy and language pedagogy requires the appreciation of experiential knowledge of a practitioner on the one hand, and the accumulation of not one but many different types of research studies, asking diverse questions and using different methodologies, on the other. Just as argued by the supporters of evidence-informed practice, ‘a wide range of information sources, empirical findings, case studies, clinical narratives and experiences are to be used in a creative and discriminating way throughout the intervention process’ (Nevo, Slonim Nevo, 2011: 3).

A very important but tiny fraction of empirical investigation needs to address the relationships between cause and effect (to estimate the so-called *efficacy*) to differentiate between the higher and lower quality teaching strategies, and especially to detect strategies which are counterproductive. And that is the narrow role of randomised controlled trials. Furthermore, randomised controlled trials verifying efficacy would never be possible until a discipline accumulates enough scholarship, from the theoretical investigation, through qualitative descriptive studies to refine enough the conceptual apparatus to satisfactorily operationalise complex ideas (like language learning outcomes) to measurable variables. Without this collective effort, even randomised controlled trials would not bring any advantageous results. There is no intention to replace multifarious scholarship with one type of methodological paradigm, and this particular paradigm answers only one type of research question. The variability of research paradigms in the context of medicine is clearly stated,

By best available external clinical evidence we mean clinically relevant research, often from the basic sciences of medicine, but especially from patient centred clinical research into the accuracy and precision of diagnostic tests (including the clinical examination), the power of prognostic markers, and the efficacy and safety of therapeutic, rehabilitative, and preventing regimens (Sackett et al., 1996: 71–72).

As we have tried to clarify that randomised controlled trials are not intended to replace other types of investigation, are they applicable to language education or other areas of education? Despite the dozens of years of successful implementation of experimental paradigms, including randomised controlled trials in the discipline of psychology, some authors argue that the nature of education excludes the use of such strict procedures in academic studies in pedagogy (e.g. Gale, 2018). One of the targets of attacks is the use of the so-called ‘control group’ and ‘placebo treatment’. The detailed methodological analysis of different variants of research paradigms from which causality may be inferred is beyond the scope of this article. We need to stress, however, that interventions in psychology and psychotherapy need to be evaluated in the most rigorous way to reduce the risk of attributing observed effects to coinciding factors and not to the actual sources of change. Control groups (receiving a different intervention, a delayed one, a placebo or no intervention) are indispensable to assess the impact of an intervention of interest. However mechanical as it may sound, this is the only way to detect well-intended but harmful strategies, examples of which we are giving in the next part of this text. Randomised

controlled trials in medicine and psychotherapy frequently lack the ‘placebo’ or ‘zero groups’ (with no treatment) due to ethical reasons; different or delayed treatment is offered in such cases. Comparisons between causal effects of two different treatments or interventions are as valid as comparisons between intervention and a placebo. Governmental, non-governmental and international bodies supervise the ethics of these practices (e.g. UK Research Integrity Office).

What is more, there is a number of excellent studies successfully applying randomised controlled trials to education, countering objections in pedagogical scholarship (e.g. for review of randomised controlled trials in education see: Connolly, Keenan, Urbanska, 2018; examples of randomised controlled trials in foreign language teaching are listed later in this article).

To summarise, the concept of evidence-informed practice supported here, which is more updated, refurbished and expanded but has its roots in evidence-based medicine, neither limits practitioners nor suppresses their expertise; it does not reduce clients to the function of their organs nor does it require fixed procedures. In education or language education, we call for empirical verification of the effects of teaching methods in varying contexts without postulating standardised testing, and we believe that randomised controlled trials are not expected to abolish other types of scholarship, while they can be successfully applied in the field of education. We see a future for mutual understanding between humanities and empirical sciences, and that pedagogy in general, and especially language pedagogy in particular, may successfully provide examples of evidence-informed practice that draws upon empirical data without restricting the practitioner’s autonomy.

## **The necessity of research-literacy and evidence-informed practice in language teaching and learning**

What would definitely enrich the practical handbooks for teachers, including language teachers, workshops for language teacher trainees, as well as the decisions of policymakers is the good quality of empirical verification of the effects that given teaching methods and techniques generate. We claim that the inclusion of effects validation not only supplements pedagogical scholarship but also reduces many risks and solves some problems in education as it is important to bear in mind that even though particular methods, approaches, or techniques are logical, popular or convincing, they do not have to be beneficial in every case or every context. The aforementioned infamous examples from medicine and psychology should be a sufficient warning, but below we focus on language education and discuss examples of recommendations that are harmful and ungrounded as well as those based on data of unsatisfactory quality.

Many speech therapists and teachers of infants and younger toddlers used to unjustifiably discourage parents from using the so-called *baby talk*. Baby talk (*motherese* or *infant-directed speech*) is simplified speech used spontaneously by adults with babies and toddlers, and it is recognizable by its highly modulated intonation. As it is different from the standard language, it was perceived as potentially damaging for early speech

development, while data proved that the early use of *motherese* is neither harmful nor indifferent but actually beneficial for early language acquisition (cf. Saint-Georges et al., 2013). This example shows that recommendations based on assumptions rather than on data may lead to harm.

In the area of writing skills, a drill characterised by unproven efficacy supposedly aimed at spelling improvement via multiple re-writings of words or sentences is of dubious utility to the students most in need of effective techniques enhancing the spelling skills, i.e. those suffering from dyslexia. Neuropsychological studies show that this type of drill relies on the so-called *implicit* or *procedural learning* (colloquially named ‘muscle memory’), and this ability is impaired in people with dyslexia (Nicolson et al., 2010), which means they cannot benefit from this form of training. This evidence should stimulate research-literate teachers to look for other techniques to accommodate pupils with special educational needs.

The examples offered above give an idea of how research-literate language teachers could enhance their practice if they drew greater attention to research data. Below, however, we will discuss the risks of informing practice in low-quality data. *Teaching and Learning Toolkit* (Education Endowment Foundation, 2020), a British synthesis of multiple evidence-based teaching strategies and approaches aggregates studies of varying quality. Regrettably, some recommendations for teaching refer to vague descriptive statistics and generalisations reminiscent of a *panacea*. For example, the analyses addressing the impact of homework are based on hundreds of studies (which could be an advantage as it should reduce the risk of accidental false findings); nevertheless, the research design of these studies is poor, variables are not defined in a satisfactory way, and conclusions are discouragingly imprecise. The typical indicator of homework practice is the average amount of time spent on homework irrespectively of the school subject (foreign language or mathematics) or the type of homework (gap-filling or creative writing). Seemingly vague are the indicators of learning outcomes which are based on the average grades obtained by pupils. Finally, the average amount of homework is correlated with the average of grades. Reports of such general scope lose utility and, what is even more dangerous, may lead to erroneous conclusions and malpractice. A medical equivalent would be a study correlating such imprecise factors as the frequency of medical consultation with such vague outcomes as overall health proving that patients who visit their doctors more frequently have worse general health than those who rarely seek medical consultation. Such erratic reasoning is explained by the Latin phrase ‘*cum hoc ergo propter hoc*’ (‘with this, therefore because of this’) and such questionable and poorly planned comparisons should not be taken into consideration. The mechanical accumulation of huge amounts of low-quality data does not lead to reliable recommendations and is equally dangerous as the lack of empirical verification. Teachers need to be research-literate to be able to differentiate between high and low-quality data, and they deserve to be empowered by good quality evidence offering real support to optimise their choices.

## **Best available evidence: solutions from psychology and psychotherapy – are they applicable to language teaching?**

The indisputable strength of academic thought in pedagogy consists in the development of sophisticated conceptualizations, elaborated analyses of the purposes and dangers within education (e.g. Biesta, 2014), and a wide array of publications offering comprehensive descriptions of teaching methods and techniques, especially in the field of language teaching, as well as classroom management strategies (e.g. Thornbury, 1999; Kelly, 2000; Harmer, 2015). However, to make responsible evidence-informed choices, teachers need to possess basic research literacy to see the difference between instructional materials offering just ‘speculations’ of outcomes and those presenting verified evidence of results concerning the use of given teaching techniques, strategies or classroom activities. Consequently, one of the crucial skills for professionals consists in the ability to assess the relative strength of different types of studies.

When teachers turn to academic texts, they may expand their understanding of interpersonal dynamics in the classroom, learning processes, or maturation of cognitive functions. They strive to use the best teaching methods and techniques but, in their professional books, they find mainly descriptions of how to implement a given solution and rarely whether it is proven to be effective. The below-presented list is meant as guidelines to evaluate different research methods. One needs to keep in mind that the credibility of the research methodology should be evaluated without losing track of the research questions addressed in the study, which are of the main interest to the practitioner, and that all empirical studies are important and provide information to revise the theoretical considerations. To put it differently, varied research methodologies are less or more adequate to answer different research questions and may be introduced at distinct stages of advancements in the state-of-the-art in a given field.

To promote teachers’ research literacy, we adjust a perspective developed by the American Psychological Association and propose guidance to evaluate the value and credibility of empirical research evidence of varying designs (based on: APA, 2006). The context of psychological consultation or psychotherapeutic intervention is not less complex than teaching; the personal variables play equally decisive roles, and yet modern practice in psychology and psychotherapy is informed through evidence to a considerable extent. Additionally, ethical standards of psychologists and psychotherapists require the practitioners to inform their clients to what extent the method used is empirically verified (see: Standard 2.04, 2.01e, 10.01b in APA, 2017).

According to the current American Psychological Association policy, the below-described spectrum of research designs deals with the ascending levels of methodological rigour and reliability. Importantly, this hierarchy does not disqualify studies of lower levels of methodological reliability. Real-life observations and case studies which may be carried out by psychotherapists or teachers as action research play a vital role in feeding not only practice but also other, more rigorous investigations. The scientific methodology requires

steady data accumulation in the field, starting from basic observational paradigms, and leading to more and more rigorous study designs. Neither preliminary less rigorous studies nor investigations using more sophisticated methodologies can be skipped. Notably, the methodological strictness of a study is just one of the factors which need to be taken into consideration besides the relevance of research questions, as many studies answer purely theoretical rather than practical problems and serve other goals than those of efficacy testing. Furthermore, the reliably proven *efficacy* of an intervention or teaching practice, defined as the prediction of probable benefits and risks is, by no means, the sole important determinant for decision-making (APA, 2002). Policymakers developing guidelines for teachers should not omit other dimensions crucial for evidence-informed practice, such as *utility*, which refers to the applicability, feasibility, or general usefulness of the intervention. One of the aspects of utility includes the generalizability of the interventions vs. their specificity, which in turn requires restraint in drawing conclusions, especially in the light of the trends related to special educational needs or universal design for education.

#### 1. Lack of evidence of effects

Research-literate practitioners need to recognise whether the proposed teaching strategy or activity is validated or not by any empirical data. Language teaching approaches, methods or techniques lacking evidence should not be excluded just on the basis of the absence of data concerning their effects. The lack of evidence does not equal evidence of the lack of positive outcomes of a given solution (cf. Chalmers, 2005) although it may be recommended to choose an alternative one, if available, with reliable evidence showing efficacy for particular needs. In the light of the lack of empirical evidence, practitioners rely on the recommendations of experts, authorities or governing bodies. Importantly,

[c]onsensus, by which we mean agreement among recognized experts in a particular area, can always add information. As the sole basis for conclusions about efficacy, consensus is more compelling than individual observation but less compelling than carefully controlled empirical evaluation (APA, 2002: 1054).

Facing the lack of evidence, research-literate teachers may take an active role in evidence collection by describing their observations in real-life settings (see below: Descriptive studies).

#### 2. Descriptive studies

Descriptive studies may use different types of methodologies, but their common characteristics consist in the lack of a causal relationship established between an intervention and its outcomes. In descriptive studies, the main aim is to notice some regularities of a heuristic value. Descriptive studies of varying levels of systematicity are the necessary steps to subsequently address the issues in a more rigorous way. An interesting example of a descriptive study may be that of Rahimi and Karkami (2015) in which Persian students evaluated the perceived effectiveness of their teachers of English as a foreign language and expressed their views of the teachers' strategies to maintain classroom discipline. The method is quantitative and correlations are shown between non-punitive strategies and perceived higher teaching effectiveness. In this study, due to its methodology, the causal link between classroom discipline strategies and teaching effects cannot be established,

but it indicates fascinating areas for future research. These results may be used as an inspiration to teachers for their reflective practice.

2.1. Observation/case study/multiple case studies (see: Merriam, 1998 – general standards for this type of study in education)

The so-called ‘clinical observation’, which sometimes appears in the form of an anecdotal short description or frequently as a more or less profound ‘case study’, is a descriptive qualitative study (though some quantitative measurements may be included) of observable effects as well as subjective perceptions of a particular type of intervention, teaching strategy, technique or activity concerning one or more subjects. Such observations and reports play a crucial role in psychology and psychotherapy and are valuable sources of innovation.

The first description of the dyslexic boy who did not respond to the proper training in reading is an excellent example of a crucial turning point in education and psychology (Morgan, 1896). Although Morgan’s original paper was just a one-page case study, it turned out to be revolutionary, since, without such an observation, the phenomenon of dyslexia would not enter the academic reflection leaving 15% of the population without the recognition of their struggle. Dyslexia had of course existed before Morgan’s first communication and what Sharples describes as ‘experiential knowledge’ in our opinion may and should be a matter of professional discourse (Sharples, 2013). While a method of observation in its nature does not constitute strong evidence, clinical/pedagogical observations may be treated as the indispensable first step leading to proper experimental research on the nature of a given phenomenon. Morgan published the observation of unexplainable difficulty in reading without indicating its nature or possible remedies. Other more rigorous methods and the accumulation of data over decades led to a better understanding of the special educational needs of pupils with dyslexia and the development of strategies promoting the best learning results (e.g. Goodwinm Ahn, 2010).

In teaching, an ‘educational observation’ may be parallel to a ‘clinical observation’ as long as it is published and disseminated. An excellent representation of a case study is provided by Han and Yao (2013), in which the authors recorded and analysed strategies used by bilingual teacher trainees who taught Chinese using English to learners of Chinese as a foreign language. They explored the use of English as the language of instruction as well as the strengths and weaknesses exhibited by the teacher trainees. The results are not conclusive in terms of teaching efficacy, but they shed valuable light on teaching practices. However, it might be useful to draw on these findings to plan a study of teaching activities using more sophisticated research methods.

2.2. Aggregated descriptive studies

More advanced types of descriptive studies may have a form of systematic case studies which aggregate experiences from interventions provided to individuals of similar characteristics. Descriptive studies of groups of well-selected participants provide valuable information on the experience of learners taking part in language education programs. Data may be obtained from interviews, questionnaires, or other measurements. It is important, however, to precisely define the intervention that is in question.

3. Preliminary efficacy testing, e.g. single-case experimental designs

‘Efficacy’ is a core term in the type of practice that is informed by evidence. It shows the extent to which an intervention provokes desirable effects on the one hand and ad-

verse side effects on the other. The term efficacy differs from ‘effectiveness’ as efficacy is measured when a given intervention is delivered under optimal but highly controlled conditions, while effectiveness is assessed in a real-world setting (Society for Prevention Research, 2004). One may talk about efficacy only if the causal link is unambiguously established and conclusions are restricted to the specific characteristics of recipients and the context; in other words, outcomes cannot be generalised to other types of individuals or contexts until proven in these conditions.

As preliminary efficacy testing is more common in clinical settings, researchers involved in language teaching to clinical groups reach for this method more easily than general language teachers. The assessment of the efficacy of the so-called enhanced conversational recast (in a natural conversation the teacher stresses the correct form of a morpheme the child tried to use) presented by Hau, Wong and Ng (2021) is a very interesting implementation of a single-case experimental design. The authors evaluate the enhanced conversational recast in four children. The first few sessions were aimed at establishing a ‘baseline’ for the learning progress of each child. Results of the following sessions, when the enhanced conversational recast was introduced, were compared to that baseline. Results were mixed, showing that only cautious generalisations, given age, sex and contextual differences, are permissible.

This type of study verifies causality and, hence, it includes a rigorous plan designed before the study begins; it contains well-defined indicators of outcomes, pre- and post-intervention measurements, and control of the intervention process (as opposed to descriptive studies, which merely register what is happening during practice). While the weakness of this paradigm consists in the small number of participants and the lack of a control group to make comparisons (subjects without any or with different intervention), its strength lies in the experimental methodology which aims to test the causal relationship between an intervention and its outcomes. This design is popular and respected in the fields of cognitive rehabilitation and behavioural interventions. It may easily be applied as action research by research-literate teachers.

4. Aggregated data from natural settings (the same as in the public health research) (see: Guest, Namey, 2015; Isaacs, 2014 – methodological issues)

This type of research uses qualitative and quantitative methodologies to investigate social, cultural, economic, and political factors that impact educational success and failure. Popular designs consist of analyses performed on registers and databases, such as PISA, school performance tables or yet national matriculation examination scores. This type of study is not designed to test the efficacy (which requires a causal link) of a given intervention but deepens understanding of more general processes influencing education and may provide data about real-life events. An excellent example of this type of study in education is the report of the Office for Standards in Education on reading results in schools using the teaching method of phonics (OFSTED, 2010). In this study, researchers selected a small number of participating schools which introduce phonics, and they compared the results of the quantitative data for standard reading measures between these selected schools and the national average.

Similarly to public health studies which provide insights into maximising community health, the analogous ‘public education research’, even in the form of big-data analyses,

is particularly useful for tracking ways to enhance the utility of teaching strategies at a political decision-making level; it may also help practitioners looking for guidelines to support their reflective practice. It is important, however, not to confuse this valuable method of research with the widespread testing for accountability which links average school results in standardised tests to increase or decrease public funds. Accountability programs do not evaluate any particular type of practice and are examples of poor-quality research (see: Wiliam, 2010).

#### 5. Process-outcome studies

Contrary to studies testing efficacy, the process-outcome studies are rather correlational, but they offer an excellent addition to good quality experimental studies as they intend to explain not just whether but also how interventions work. For example, in psychotherapy, it is crucial to know not only which psychotherapeutic approach is the most effective for specific needs but also which elements of a complex approach play a decisive role. In this type of study, key variables representing intervention elements of the process are pre-selected, and their real-life application is measured (usually using questionnaires or observational scales) and correlated with the outcomes of the intervention (see: Llewelyn et al., 2016). There are good examples of the application of this approach to identifying mechanisms of success in education (e.g. an outstanding study of the impact of different elements of classroom management on creating a secure learning environment: Egeberg et al., 2016).

#### 6. Studies of interventions delivered in naturalistic settings (effectiveness research)

Effectiveness studies are high-quality studies verifying the causal relationship between a given practice and its outcomes. Once there are at least two different teaching strategies with proven efficacy (verified in the experimental paradigm), they may and should be compared in real-life settings. Effectiveness studies are almost equally rigorous as efficacy studies. This type of research requires two groups of participants and random attribution of two different interventions. Contrary to the prudently selected participants for efficacy studies, in this model, all real-life factors influence and modulate outcomes. This type of study provides the so-called ecological validity of interventions. One may even argue that this type of data offers a final test of the utility of interventions. Effectiveness studies seem to be ideal to verify the effects of teaching methods.

#### 7. High standard efficacy testing (compare: 3. Preliminary efficacy testing, above)

Research methodology maximising the credibility of results needs to apply experimental design (establishing a causal link) to test a hypothesis of a relationship between chosen variables. To formulate useful hypotheses, select adequate variables, and calibrate appropriately the measures, one needs to already have a solid understanding of a studied area from previous studies (case studies, databases analyses, etc.). Hypotheses are tested with the application of advanced mathematical-statistical methods to minimise the risks of false conclusions.

##### 7.1. Quasi-experimental design study

Neither psychotherapy nor teaching can realistically count on the wide application of randomised controlled trials (see below) due to multiple practical and ethical reasons. However, there are other only slightly less reliable procedures for drawing causal inferences



about the effects of interventions. A quasi-experimental design study establishes a causal relationship between independent (e.g. teaching technique) and dependent variables (e.g. teaching result) even though it lacks randomised samples; in other words, students are not included in groups via random selection, but they already belong to a group that is selected to become an experimental or control group. For example, we may test outcomes of a given strategy in typical pupils and compare them with a special educational needs group; this would be an important validation of the intervention for specific types of students. Quasi-experimental design studies give a very strong and reliable source of evidence.

### 7.2. Randomised controlled trials (RCT)

This type of experimental design has the greatest power to test hypotheses as it represents the most rigorous experimental design for testing positive and negative effects caused by a treatment. In evidence-based medicine, no medical procedure is approved before checking it via this strict evaluation to minimise patients' harm and maximise their benefits. RCT design is widely accepted in psychology and psychotherapy (Cook, Schwartz, Kaslow, 2017). There are numerous excellent examples of RCT studies in education and language education, such as the study on the narrative development of bilingual children (Uchikoshi, 2005) or the study on the use of WhatsApp to enhance spontaneous communication in EFL (Minalla, 2018). In psychology and teaching, the challenges of precise measurements of pre- and post-intervention functioning are demanding but not insurmountable. Nevertheless, the RCT attracts vigorous opposition in some educational circles as has been previously discussed (see: Gale, 2018).

8. Synthesis of multiple evidence (see: Beretvas, 2005 – on challenges in summarising empirical studies)

Even the most reliable research methodology may lead to false conclusions. To exclude accidental results of a single study, the scientific methodology requires effects to be replicated in different contexts, preferably by different researchers and on different subjects. This sometimes leads to contradictory conclusions from multiple papers and what is needed is a strategy to evaluate and synthesise numerous original findings.

8.1. The basic method of integration of different studies is a review of papers, in which the author provides evidence to answer research questions through qualitative analyses and discussion of relevant data.

8.2. A more reliable method is the so-called systematic review which is based on the strict selection of quality papers and provides an exhaustive summary as well as an in-depth analysis of current evidence in the field (as, for example, the previously cited systematic review of standardised testing on students, Harlen, Deakin-Crick, 2002).

8.3. The best and the most appreciated method to combine results from multiple studies consists in the so-called 'meta-analysis'. Similarly to the systematic review, it offers a list of selected original papers adhering to strict criteria; it differs from the systematic review, which is qualitative in nature, in that the meta-analysis provides quantitative results through the statistical estimation of the size of effects from multiple original papers. What is also essential to underscore is that it matters what quality of design is represented by the original papers analysed. Accordingly, the most reliable meta-analyses are those which synthesise randomised controlled trials. The application of meta-analysis methodology to calculate

the size of effects of low-quality primary data does not result in reliable conclusions as we have already discussed it in the previous section of this paper (the example of homework effects in education, Education Endowment Foundation, 2020).

The study by Adesope et al. (2011) may be given as an example of a meta-analysis of good-quality studies, as the selection criteria were to include only papers describing results of experimental or semi-experimental designs. This meta-analysis offers reliable information on ‘what works’ in the context of teaching literacy to EFL immigrant students. This type of data permits teachers, educators and policymakers to inform their decisions on the relative strengths of different pedagogical approaches.

## Conclusions

In applied disciplines like medicine, psychology and psychotherapy, we may observe a shift from practice mainly inferred from theoretical assumptions to basing it more and more on empirically verified effects. The introduction of this paradigm to teaching faces many obstacles voiced by scholars and practitioners who seem rather cautious about implementing methodological routines of empirical sciences in humanities. In evidence-informed teaching, insights originating from each and every type of research design, serving different objectives, play an important role. Qualitative research methodology, drawing upon experiential knowledge, in particular, may be applied by research-literate teachers in their action research to enrich the common reservoir of knowledge regarding the benefits of given classroom activities targeting different teaching contexts. Only after multiple qualitative studies have been conducted and published by practitioners or researchers, academics will be sufficiently informed to proceed and plan other types of studies, including those establishing causal relationships, to provide data of greater reliability and validity.

Studies show that teachers who perceive themselves as research-literate more readily turn to evidence-informed practice than those who feel less secure in their skills to evaluate research quality (Georgiou et al., 2020). Hence, research literacy may translate into the selection of more effective methods which will benefit the students. Additionally, in our view, evidence-informed practice has the potential to restore teachers’ agency as their role is not just to deliver given content in accordance with methodological preparation but to be aware of unique teaching contexts and responsibly choose the best strategies in the light of personal experiential knowledge and solid and verified empirical data. This would be a student-centred approach based on the best available evidence. In this, teachers may follow psychologists and psychotherapists, who already insist on grounding their practice not just in theory or official recommendations but also in the reliable evidence of benefits and risks.

Scholars in education may prefer inspiring general questions and discussions instead of laborious testing of the effects that different teaching approaches, methods, and techniques lead to in different students and varying contexts. But, like psychotherapists, teachers deserve precise data on risks and promises of available teaching methods, strategies, techniques or classroom activities. In our view, an evidence-informed approach requires

a bidirectional transfer of information between academics and teachers, in which the experiential knowledge and action research offered by teachers and data obtained via rigorous methodology provided by scholars would represent a reciprocal relationship. To make it happen, teacher educators need to provide not just basic training in the potential uses of different research methods but to facilitate teachers' access to academic publications, explaining their utility for various aspects of teaching practice. As Evans et al. (2017) postulate, strong collaborative links between universities, research centres and school teachers must be established. Research literacy is not a theoretical concept but a practical skill; hence, such a collaboration plays a key role in 'supporting the sustainability of research and in enabling teachers to connect their own practice with the broader body of research knowledge. Teacher and pupil ownership of research is crucial in developing research-integrated learning' (ibidem: 404).

## References

- Adesope O.O., Lavin T., Thompson T., Ungerleider C. (2011), *Pedagogical Strategies for Teaching Literacy to ESL Immigrant Students: A Meta-Analysis*, "British Journal of Educational Psychology", 81(4), p. 629–653.
- Alderson J.C., Wall D. (1993), *Does washback exist?*, "Appl. Linguist.", 14, p. 115–129.
- Amir A., Mandler D., Hauptman S., Gorev D. (2017), *Discomfort as a means of pre-service teachers' professional development – an action research as part of the 'Research Literacy' course*, "European Journal of Teacher Education", 40(2), p. 231–245.
- APA (2002), *Criteria for Evaluating Treatment Guidelines*, "American Psychologist", 57(12), p. 1052–1059.
- APA (2006), *Evidence-based practice in psychology*, "American Psychologist", 61(4), p. 271–285.
- APA (2017), *Ethical Principles of Psychologist and Code of Conduct*, <https://www.apa.org/ethics/code>
- Bailey K.M. (1996), *Working for washback: a review of the washback concept in language*, "Lang. Test.", 13, p. 257–279.
- Barnes M. (2016), *The washback of the TOEFL IBT in Vietnam*, "Austral. J. Teacher Educ.", 41(7), p. 157–174.
- Becher A., Lefstein A. (2020), *Teaching as a Clinical Profession: Adapting the Medical Model*, "Journal of Teacher Education", p. 1–2.
- Beretvas S.N. (2005), *Methodological Challenges Encountered in Summarizing Evidence-Based Practice*, "School Psychology Quarterly", 20(4), p. 498–503.
- Biesta G. (2014), *The beautiful risk of education*, London: Routledge.
- Burch D. (2013), *Taken in Vein*, "Natural History", 121(5), p. 10–13.
- Center on Education Policy (2019, December), *A Stronger Future for Evidence-Based School Improvement in ESSA*, <https://files.eric.ed.gov/fulltext/ED602940.pdf>

Chalmers I. (2005), *If evidence-informed policy works in practice, does it matter if it doesn't work in theory?*, "Evidence and Policy", 1(2), p. 227–242.

Civil Rights Project (2019), *The Striking Outlier: The Persistent, Painful and Problematic Practice of Corporal Punishment in Schools*, [https://civilrightsproject.ucla.edu/research/k-12-education/school-discipline/the-striking-outlier-the-persistent-painful-and-problematic-practice-of-corporal-punishment-in-schools/COM\\_Corporal-Punishment\\_FINAL-Web.0.pdf](https://civilrightsproject.ucla.edu/research/k-12-education/school-discipline/the-striking-outlier-the-persistent-painful-and-problematic-practice-of-corporal-punishment-in-schools/COM_Corporal-Punishment_FINAL-Web.0.pdf)

Connolly P., Keenan C., Urbanska K. (2018), *The trials of evidence-based practice in education: A systematic review of randomised controlled trials in education research 1980–2016*, "Educational Research", 60(3), p. 276–291.

Cook S.C., Schwartz A.C., Kaslow N.J. (2017), *Evidence-Based Psychotherapy: Advantages and Challenges*, "Neurotherapeutics: the journal of the American Society for Experimental NeuroTherapeutics", 14(3), p. 537–545, <https://doi.org/10.1007/s13311-017-0549-4>

Cooper A., Klinger D.A., McAdie P. (2017), *What do teachers need? An exploration of evidence-informed practice for classroom assessment in Ontario*, "Educational Research", 59:2, p. 190–208.

Education Endowment Foundation (2020), *Teaching and Learning Toolkit*, <https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/>

Egeberg H.M., McConney A., Price A.E. (2016), *Classroom Management and National Professional Standards for Teachers: A Review of the Literature on Theory and Practice*, "Australian Journal of Teacher Education", 41(7), p. 1–18.

Evans C., Waring M., Christodoulou A. (2017), *Building teachers' research literacy: integrating practice and research*, "Research Papers in Education", 32(4), p. 403–423.

Gale T. (2018), "What's Not to like about RCTs in Education?" [in:] A. Childs, I. Menter (Eds.), *Mobilising Teacher Researchers: Challenging Educational Inequality* (p. 207–223), Abingdon: Routledge, <http://eprints.gla.ac.uk/158072/1/158072.pdf>

Georgiou D., Mok S.Y., Fischer F., Vermunt J.D., Seidel T. (2020), *Evidence-Based Practice in Teacher Education: The Mediating Role of Self-Efficacy Beliefs and Practical Knowledge*, "Frontiers in Education", 5:559192, <https://doi.org/10.3389/feduc.2020.559192>

Goodwin A.P., Ahn S. (2010), *A meta-analysis of morphological interventions: effects on literacy achievement of children with literacy difficulties. A meta-analysis of morphological interventions: effects on literacy achievement of children with literacy difficulties*, "Annals of Dyslexia", 60, p. 183–208.

Guest G., Namey E.E. (2015), *Public Health Research Methods*, London: SAGE Publications.

Han J., Yao J. (2013), *A Case Study of Bilingual Student-Teachers' Classroom English: Applying the Education-Linguistic Model*, "Australian Journal of Teacher Education", 38(2), p. 118–131.

Harlen W., Deakin-Crick R. (2002), *A systematic review of the impact of summative assessment and tests on students' motivation for learning*, [in:] EPPi-Centre (Ed.), *Research evidence in education library*, London: Institute of Education Social Science Research Unit.

Harmer J. (2015), *The Practice of English Language Teaching*, London: Pearson.

Hau F.F.-W., Wong A.M.-Y., Ng M.W.-Y. (2021), *Does Enhanced Conversational Recast Promote the Learning of Grammatical Morphemes in Cantonese-Speaking Preschool Children? Answers from a Single-Case Experimental Study*, "Child Language Teaching and Therapy", 37(1), p. 43–62.

- Isaacs A.N. (2014), *An overview of qualitative research methodology for public health researchers*, "International Journal of Medicine and Public Health", 4(4), p. 318–323.
- Kanner L. (1943), *Autistic disturbances of affective contact*, "Nervous Child", 2, p. 217–250.
- Kelly G. (2000), *How to teach pronunciation*, Harlow: Pearson Education Limited.
- Lagay F. (2002), *The legacy of humoral medicine*, "AMA Journal of Ethics", 4(7), p. 206–208.
- Lindhiem O., Bennet C.B., Trentacosta C.J., McLearn C. (2014), *Client Preferences Affect Treatment Satisfaction, Completion, and Clinical Outcome: A Meta-Analysis*, "British Journal of General Practice", 64(628), p. 506–517.
- Llewelyn S., Macdonald J., Aafjes-van Doorn K. (2016), *Process–outcome studies*, [in:] J.C. Norcross, G.R. VandenBos, D.K. Freedheim, B.O. Olatunji (Eds.), *APA handbooks in psychology. APA handbook of clinical psychology: Theory and research*, American Psychological Association, p. 451–463.
- Mausethagen S., Raaen F.D. (2017), *To jump the wave or not: teachers' perceptions of research evidence in education*, "Teacher Development", 21(3), p. 445–461.
- McKnight L., Morgan A. (2019, March 25), *The problem with using scientific evidence in education (why teachers should stop trying to be more like doctors)*, "Australian Association for Research in Education", <https://www.aare.edu.au/blog/?p=3874>
- Merriam S.B. (1998), *Qualitative Research and Case Study Applications in Education: Revised and Expanded from Case Study Research in Education*, (2nd edition), San Francisco: Jossey-Bass.
- Minalla A.A. (2018), *The Effect of WhatsApp Chat Group in Enhancing EFL Learners' Verbal Interaction outside Classroom Contexts*, "English Language Teaching", 11(3), p. 1–7.
- Morgan W.P. (1896), *A Case of Congenital Word Blindness*, "The British Medical Journal", 2(1871), p. 1378.
- Nelson J., Campbell C. (2017), *Evidence-informed practice in education: meanings and applications*, "Educational Research", 59:2, p. 127–135.
- Neufeld K. (1990), *Preparing Future Teachers as Researchers*, "Education", 110(3), p. 345–351.
- Nevo I., Slonim-Nevo V. (2011), *The Myth of Evidence-Based Practice: Towards Evidence-Informed Practice*, "The British Journal of Social Work", 41(6), p. 1176–1197, <https://doi.org/10.1093/bjsw/bcq149>
- Ng B-Y. (1999), *Hysteria: A cross-cultural comparison of its origins and history*, "History of Psychiatry", 10(39), Special Section: Transcultural Psychiatry, p. 287–301.
- Nicolson R.I., Fawcett A.J., Brookes R.L., Needle J. (2010), *Procedural learning and dyslexia*, "Dyslexia", 16(3), p. 194–212.
- Nordenstrom J. (2007), *Evidence-Based Medicine: In Sherlock Holmes' Footsteps*, Boston: Blackwell Publishing.
- OFSTED, The Office for Standards in Education, Children's Services and Skills (2010), *Reading by six: How the best schools do it*, [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/379093/Reading\\_20by\\_20six.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/379093/Reading_20by_20six.pdf)
- Preference Collaborative Review Group (2009), *Patients' preferences within randomised trials: systematic review and patient level meta-analysis*, "BMJ: British Medical Journal", 338, p. 85–88.

Rahimi M., Karkami F.H. (2015), *The Role of Teachers' Classroom Discipline in Their Teaching Effectiveness and Students' Language Learning Motivation and Achievement: A Path Method*, "Iranian Journal of Language Teaching Research", 3(1), p. 57–82.

ResearchEd (n.d.), *What is researchED?*, <https://researched.org.uk/>

Sackett D., Rosenberg W.C., Gray J.A.M., Haynes R.B., Richardson W.S. (1996), *Evidence based medicine: what it is and what it isn't*, "BMJ", 312, p. 312–371.

Saint-Georges C., Chetouani M., Cassel R., Apicella F., Mahdhaoui A., Muratori F., Laznik M., Cohen D. (2013), *Motherese in Interaction: At the Cross-Road of Emotion and Cognition? (A Systematic Review)*, "PLOS ONE", 8(10), p. 1–17.

Sari M. (2006), *Teacher as a Researcher: Evaluation of Teachers' Perceptions on Scientific Research*, "Educational Sciences: Theory & Practice", 6(3), p. 880–887.

Sharples J. (2013), *Evidence for the Frontline*, London: Alliance for Useful Evidence, <https://apo.org.au/sites/default/files/resource-files/2013-06/apo-nid34800.pdf>

Society for Prevention Research (2004), *Standards of Evidence Criteria for Efficacy, Effectiveness and Dissemination*, <https://www.preventionresearch.org/StandardsOfEvidencebook.pdf>

Sousa A.C. (2011), *From Refrigerator Mothers to Warrior-Heroes: The Cultural Identity Transformation of Mothers Raising Children with Intellectual Disabilities*, "Symbolic Interaction", 34(2), p. 220–243.

Thornbury S. (1999), *How to Teach Grammar*, Harlow: Longman.

Tuñón J. (2002), *Creating a Research Literacy Course for Education Doctoral Students*, "Journal of Library Administration", 37(3–4), p. 515–527.

Uchikoshi Y. (2005), *Narrative development in bilingual kindergarteners: Can Arthur help?*, "Developmental Psychology", 41(3), p. 464–478.

Van Zyl J., Nel K., Govender S. (2017), *Reparative sexual orientation therapy effects on gay sexual identities*, "Journal of Psychology in Africa", 27(2), p. 191–197.

Waring M., Evans C. (2015), *Understanding Pedagogy: Developing a Critical Approach to Teaching and Learning*, Abingdon: Routledge.

William D. (2010), *Standardized Testing and School Accountability*, "Educational Psychologist", 45(2), p. 107–122.

Woodbury M.G., Kuhnke J.L. (2014), *Evidence-based Practice vs. Evidence-informed Practice: What's the Difference?*, "Wound Care Cana", 12(1), p. 18–21.

## Streszczenie

### Rozwijanie umiejętności badawczych u nauczycieli języków: zobrazowanie trudności i wyznaczenie kierunków praktyki wspieranej badaniami

Celem tekstu jest pokazanie potencjalnych korzyści płynących z rozwijania umiejętności badawczych nauczycieli w ogóle i nauczycieli języków w szczególności. Sądzimy, że nauczyciele, którzy nabyli umiejętności badawcze, są w stanie efektywnie czerpać z badań empirycznych przy podejmowaniu decyzji w klasie, co może zoptymalizować korzyści uczniów. Podkreślamy fakt, że odwołując się do swojego doświadczenia zawodowego, nauczyciele mogą również pełnić rolę aktywnych badaczy dzielących się swoimi obserwacjami i weryfikujących skuteczność poszczególnych technik nauczania lub działań w różnych kontekstach

i dla indywidualnych potrzeb edukacyjnych. Próbując wyjaśnić podstawowe nieporozumienia, dostrzegamy potrzebę praktyki nauczania wspieranej badaniami empirycznymi, która wywodzi się z podejścia „opartego na dowodach”, popularnego w medycynie, psychologii i psychoterapii. W ostatniej części tekstu pokazujemy przydatność różnych metod badań empirycznych do odpowiedzi na różnorodne pytania badawcze, podając jednocześnie przykłady wysokiej jakości badań z zakresu edukacji językowej. Podsumowując, apelujemy o ściślejszą współpracę i dwukierunkową wymianę know-how między nauczycielami a naukowcami.

**Słowa kluczowe:** umiejętność prowadzenia badań, praktyka nauczania wspierana badaniami empirycznymi, praktyka oparta na dowodach, badania skuteczności w edukacji, nauczyciele języków, kształcenie językowe