

# An application of programmatic assessment for learning (PAL) system for general practice training

## Abstract

**Aim:** Programmatic assessment for learning (PAL) is becoming more and more popular as a concept but its implementation is not without problems. In this paper we describe the design principles behind a PAL program in a general practice training context.

**Design principles:** The PAL program was designed to optimise the meaningfulness of assessment information for the registrar and to make him/her use that information to self regulate their learning. The main principles in the program were cognitivist and transformative. The main cognitive principles we used were fostering the understanding of deep structures and stimulating transfer by making registrars constantly connect practice experiences with background knowledge. Ericsson's deliberate practice approach was built in with regard to the provision of feedback combined with Pintrich's model of self regulation. Mezirow's transformative learning and insights from social network theory on collaborative learning were used to support the registrars in their development to become GP professionals. Finally the principal of test enhanced learning was optimised.

**Epilogue:** We have provided this example explain the design decisions behind our program, but not want to present our program as the solution to any given situation.

**Keywords:** Assessment, Programmatic assessment, General practice training

Lambert Schuwirth<sup>1,2,3,4</sup>

Nyoli Valentine<sup>5</sup>

Paul Dilena<sup>5</sup>

1 Flinders Universität, Adelaide, Australia

2 Universität Maastricht, Maastricht, The Netherlands

3 Chang Gung Universität, Taiwan

4 Uniformed Services University, USA

5 Sturt Fleurieu GP, GPEX, Australia

## Background

Programmatic assessment for learning (PAL) is rapidly gaining popularity around the world [1], [2], [3]. This is quite surprising, because the concept is fundamentally different to what has been the custom in assessment in the past. Traditionally, assessment focused almost entirely on determining whether a student had learnt enough to prevent not-yet-competent students from progressing. Where there was an influence of assessment on student learning, it was mainly used from a behaviourist viewpoint; that is, passing the assessment was the proverbial "carrot" and failing was the "stick". However, the way assessment influences student learning is far more complex than this. Cilliers et al. for example showed the myriads of interactions between sources of impact on learning, the mechanisms by which learning is impacted and the possible consequences [4], [5]. This understanding of the relationship between assessment and student learning behaviour is important in using assessment specifically to direct student learning in a more meaningful way. This is the main purpose of assessment for learning [6], [7].

In order for assessment to drive student learning in a more meaningful way, the student has to be enabled to obtain meaningful information about their own performance and use this for their future learning. This is where

the programmatic aspect of PAL comes in [8]. Just providing a score and a pass-fail decision, for example, does not inform the student sufficiently about their strengths and weaknesses. So, it does not help them in formulating more specific and concrete learning plans. Also in many traditional assessment programs, combining information is based on the format of the assessment; in an OSCE, for example, performance on an abdominal examination station is combined with performance on a knee examination. In PAL, information is combined across different assessment methods to make it more meaningful. This principle is perhaps best illustrated using a clinical example [9]. When we combine our patient's complains about fatigue, thirst and frequent urination with physical examination findings such as poorly healing wounds and absent peripheral arterial pulsations, and with the numerical value of 32 mmol per litre for the blood glucose it easily adds up to "diabetes mellitus". We would not consider telling a patient that their glucose level is far too high but fortunately their potassium level is far too low so on average they are okay, which would be the clinical equivalent of combining the performance on an abdominal examination patient with the performance on a knee examination station. In programmatic assessment we therefore aim to combine information across assessment methods in the same way to diagnose "dyscompet-

ence". Of course the disease "dyscompetence" does not exist but we use the term merely to illustrate principle.

An additional feature of programmatic assessment is the concept of proportionality. This means that the stakes of the decisions made about a student or learner have to be proportional to the credibility of the underlying information. So, single observations or single assessments can be used for feedback but not for high-stakes decisions. In programmatic assessment, all single observations or assessments are collected and collated over time until sufficient information is available to make a high-stakes decision [10]. This again, is quite similar to daily clinical practice; we are comfortable with making a simple diagnosis – for example an upper respiratory tract infection – on the basis of little information but for a high-stakes diagnosis – such as a malignancy – we want to rely on multiple sources of diagnostic information (lab values, imaging, pathology, et cetera).

So logically, programmatic assessment is a longitudinal approach to assessment in which the outcomes of many formal and informal assessments are collected on a continual basis, for example in a portfolio. Typically, the learner and a staff member – often called a mentor or coach – meet at regular intervals to discuss the learner's progress and their concrete learning goals. And at the end of the study phase, all information is used to decide whether the learner is allowed to progress or not. Importantly, the learner makes an analysis based on all the information available to him or her, and formulates concrete learning goals before meeting with their mentor or coach. As such, the coach is able to give a prognosis during the phase of the most likely outcome. The concept of programmatic assessment has been described in various publications [1], [7], [8], [10].

Typically, when the concept of PAL is explained – especially to healthcare providers – it feels intuitively right and people are willing to accept the concept, but implementation is not at all easy. There are several reasons for this. The first and probably most important reason, is the fact that it is a fundamental change in thinking about the role of assessment. Fundamental changes to any discipline just need time to find their way from theory to practice. A conceptually different way of thinking about education, such as problem-based learning, has taken many years and even decades before it has become widely accepted. A second reason concerns the logistical changes needed for the implementation. In PAL the whole assessment program is explicit and therefore the associated investment in time and costs are overt, whereas in many traditional assessment programs costs are generally more covert. That makes a cost comparison quite difficult and easily result in a negative perception towards PAL. The third reason lies most likely in what Vosniadou calls naïve frameworks or naïve theories [11]. Through our experiences in the world we develop our own views on how the world works and it is very difficult to change these beliefs. They can be complemented by more formal theories, but they never really disappear. This also happens with education. Our views on what education is and how it should

be organised have been shaped by our lengthy experience as learners ourselves, and although they can be complemented by formal training – such as staff development or teacher training – they never completely disappear. So when an attempt is made to implement PAL the naïve beliefs still continue to influence the various detailed design decisions.

Some of the more intuitive approaches to deal with beliefs and the management of such a change concern careful identification of stakeholders and their roles and careful communication with them. This communication needs to be open and continuous but also agile in that it relates to different stakeholders with different arguments and explanations; ranging from evidence from research to rhetorical conviction. In addition, one of the factors that can help in this process is a description of implementations somewhere else; as a demonstration that the concept can be translated to an actual practice. The purpose of this paper is to provide such a proof of concept. In the Consensus statement and recommendations from the 2010 Ottawa conference [12] the ideographic description or educational case report is seen as an important type of research as long as it connects the described practice with the underlying theoretical concepts so that it allows the readers to understand the design decisions and adapt them to their own context. In medical education, this 'adaptability' is considered more helpful than mere replicability of findings [13]. This paper therefore presents an educational case study.

## Context

GP365 is a general practice training program in South Australia and Western Australia, which has been developed by Sturt Fleurieu Education and Training in collaboration with the Flinders University Prideaux Centre for Research in Health Professions Education. It is a one-year curriculum in the context of a three or four-year training program and it is run for all GP registrars in South Australia and Western Australia. During this year, GP registrars (residents) follow the GP365 program which supports them through linking their practice experiences to background knowledge, skills and understanding. During this year, GP365 provides registrars with background reading material, assignments, a supervisor, a medical educator and a peer group to work with. On a continual basis registrars receive feedback from their supervisors, from medical educators and from their peers. In addition, they are given formative tests on relevant knowledge and application of knowledge. During their training registrars build a portfolio which will eventually contain evidence of all feedback from: directly observed patient consultations, reviewed videoed consultations, critical case analysis write-ups, a clinical audit they have performed, professionalism, activities in their peer group, multi-source feedback, mid and end term assessments, the results of their formative tests with their own analyses and written feedback from their medical educator.

In total, this program may not seem innovative or different to what is done in many postgraduate training contexts, but there are differences which we will explain below.

## Design principles

The most important design principle behind GP365 is the problem of transfer and understanding of the so-called “deep structure” [14]. Registrars, during their training, see a huge variety of individual patient cases, but to become an expert it is important to understand similarities and differences between these cases; thus to build transfer [15]. The literature describes the importance of decontextualizing, understanding first principles and re-contextualising (applying those principles in another case) for the development of transfer and expertise [15], [16]. Therefore, in GP365 the assessment seeks to support the registrar in meaning making of these individual experiences, for example by requiring them to relate basic medical sciences and background clinical knowledge to their individual patient experiences. This is typically what the critical case analysis write-ups focus on. The registrar chooses a patient for their critical case analysis write up, but has to be able to explain why they see that particular patient as most relevant to their own learning. The registrar also defines their own concrete learning goals and then studies the necessary background information to obtain a complete understanding of the clinical case and its management. As “evidence” of this learning the registrar produces three case based multiple-choice questions backed up by a literature reference. He or she then submits the case write up and questions. The medical educator reads the clinical case write up and provides ample feedback which the registrar has to implement in a revised version of the case write up. As such this is an activity that requires the registrar to make optimal meaning of what they have experienced during their practice but with evidence of the related learning. This evidence is reviewed and will be a mandatory part of their portfolio – as are all sorts of evidence – and contribute to the final decision.

The second design principle is the aggregation of information across different assessment parts. An example of such connection starts with the critical case analysis write-ups. As described above, the registrar receives feedback on all their write-ups which they have to implement and revise. This is an application of Ericsson’s principle of deliberate practice [17], [18]. The multiple-choice questions that each registrar generates are collected into an item bank. From this bank, periodically, progress tests are constructed and presented to the registrars, which they can sit during a predefined time window using the electronic learning system. The test items are then released and the registrars are required to critique at least three questions; preferably those questions they find most contentious. When they critique questions, they have to provide copies of the relevant scientific literature supporting their critique. The idea behind this is to optimise

the influence of test enhanced learning by asking the registrars to critically review the items and their own responses [19] and has been used in progress testing in various settings [20], [21]. These critiques are discussed in a peer group meeting. After this exchange of critiques between the members of the group, the group is expected to produce a consensus of the questions that they find most contentious and with a summary of their critique. Only then will the registrars receive their scores on the test, which they can then analyse and use in their portfolio. The design principle behind the group meetings is to foster the development of informal peer networks. Registrars may be practising in remote areas, and many have limited or no colleagues their own age or experience in their practice which they feel they can relate too. The literature shows that the possession of informal networks is important for receiving information and learning [22], [23]. A further included principle from transformative learning theory relates to making registrars aware that not all which is written is necessarily unambiguously true, and that tolerance for uncertainty is part of practice [24]. This is an illustration of how the assessment programme leads the registrars to integrate information from the assessment in a more meaningful way across instruments, so as to optimise the ‘constructivist’ drivers of the assessment on learning.

A third principle is an increase in self responsibility for learning. Generally, after graduation learners are expected to be able to take control of their own learning and assessment. Unfortunately, this is not always the case. One of the problems with CME, for example, is that people tend to follow courses in those areas there are already good at [25]. Assessment for learning should actually equip learners with the ability to analyse their strengths and weaknesses, to translate these into specific learning goals and to actually make that learning occur. For most people, self-regulated learning does not come naturally and it requires development and guidance. Paul Pintrich’s model is helpful as it distinguishes activities such as: “forethought, planning and activation”, “monitoring”, “control” and “reaction and reflection”; each of which requires learners to manage their cognition, motivation, behaviour and context [26]. Of course, having regular meetings and requiring the registrars to constantly analyse their own progress, strengths and weaknesses, and by asking them to formulate concrete do-able learning goals, the elements of “forethought, planning and activation”, “monitoring”, “control”, and “reaction and reflection” are built into the assessment system. Registrars who fail to undertake or do not complete these self-regulation activities in a comprehensive manner, are required to re-do the activities and will eventually not be allowed to progress. By providing the registrars with feedback, a peer group, a supervisor and a dedicated medical educator, the program supports the registrars’ motivation and behaviour, and by providing them with ample information through the assessment it supports their cognition. Their relationships with their supervisor, medical educator and their peer groups provide support around learning how

to navigate their current and future complex context. Registrars who do not progress as desired will have to follow remediation. But, they themselves will have to take action and design their own remediation, of course with the support and permission of their supervisor and medical educator. As such, they will have to take full responsibility for their own learning, like they will have to do after graduation.

A fourth principle is the longitudinality of the programme. All the information about the registrar's performance is collected in a portfolio which is discussed periodically with their supervisor and/or medical educator. Initially, when the information in the portfolio is still 'thin' mostly more formative feedback is given, when the information becomes richer serious suggestions for intervention and remedial activities are given and the final decision whether or not the registrar is ready to progress to the next phase is always based on the full and rich information.

A final design principle is adaptation of the program to the local context. It is unlikely that any approach in medical education which works well in one country or context could be simply applied in another country. In order for education to be successful it has to link with expectations and cultural determine of its environment. In GP365 it was therefore important to realise that Australia is a vast country with many remote areas. GPs play an important role in the Australian healthcare context and almost always they are the first port of call; moreover, they are often the only port of call. This puts a specific stress on educational programs to educate GPs who are optimally equipped to work individually and safely. Therefore, elements such as self-directed learning, accountability, dealing with uncertainty and tolerance for ambiguity features so prominently in the program, for example like the peer group meetings – the so-called mini releases –, the ongoing feedback on professionalism and the frequent meetings with experienced supervisors and medical educators. This is woven into the program in all kinds of aspects to ensure that the program meets the Colleges' requirements of "safe independent practitioner" and "college exam readiness".

## Epilogue

The program has now been running for two years and it is too soon to provide extensive information as to its success with respect to the quality of graduates. Early results do suggest that the pass rates on fellowship exams are far above the national average but before a clear causal relationship with the PAL approach can be argued for more data are needed. That was not the intent of this paper and although there are anecdotal indications that programmatic assessment for learning is effective, the whole approach is still too young to have produced a sufficiently credible body of evidence. So the jury is still out. Moreover, we feel that any success in our context would not mean that the exact same program would have

the same success in another context. That's why we have focused on providing examples on how the concept of programmatic assessment for learning has influenced design decisions around GP365 assessment program. We hope that our explanation of the underlying principles and the description of how they have influenced our design decisions are helpful for anyone who is considering changing or rebuilding programmatic assessment for learning program.

## Competing interests

The authors declare that they have no competing interests.

## References

1. Van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39(3):309-317. DOI: 10.1111/j.1365-2929.2005.02094.x
2. Fishlede A, Henson L, Hull A. Cleveland Clinic Lerner College of Medicine: An innovative approach to medical education and the training of physician investigators. *Acad Med*. 2007;82(4):390-396. DOI: 10.1097/ACM.0b013e318033364e
3. Schuwirth LW, Ward H, Heeneman S. Assessment for Learning. In: Higgs J, Baldry-Currens J, Jensen G, Letts W, Sheehan D (Hrsg). *Realising Exemplary Practice-based Education*. Rotterdam: Sense Publishers; 2013. S.143-150. DOI: 10.1007/978-94-6209-188-7\_15
4. Cilliers FJ, Schuwirth LW, Herman N, Adendorff HJ, Van der Vleuten CP. A model of the pre-assessment learning effects of summative assessment in medical education. *Adv Health Sci Educ Theory Pract*. 2012;17(1):39-53. DOI: 10.1007/s10459-011-9292-5
5. Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, Van der Vleuten CP. The mechanisms of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract*. 2010;15(5):695-715. DOI: 10.1007/s10459-010-9232-9
6. Shepard L. The role of assessment in a learning culture. *Educ Res*. 2009;29(7):4-14. DOI: 10.3102/0013189X029007004
7. Schuwirth L. From assessment of learning to assessment for learning. Huntington WV: IAMSE; 2107. S.170. Zugänglich unter/available from: <http://www.iamse.org>
8. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach*. 2011;33(6):478-485. DOI: 10.3109/0142159X.2011.565828
9. Schuwirth LW, Van der Vleuten CP, Durning SJ. What programmatic assessment for learning in medical education can learn from healthcare. *Perspect Med Educ*. 2017;6(4):211-215. DOI: 10.1007/s40037-017-0345-1
10. Van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol*. 2010;24(6):703-719. DOI: 10.1016/j.bpobgyn.2010.04.001
11. Vosniadou S. Capturing and modeling the process of conceptual change. *Learn Instruct*. 1994;4:45-69. DOI: 10.1016/0959-4752(94)90018-3



12. Schuwirth LW, Colliver J, Gruppen L, Mennin S, Kreiter C, Mennin S, Onishi H, Pangaro L, Ringsted C, Swanson D, Van der Vleuten C, Wagner-Menghin M. Research in assessment: Consensus statement and recommendation from the Ottawa 2010 conference. *Med Teach*. 2011;33(3):224-233. DOI: 10.3109/0142159X.2011.551558
13. Cianciolo AT, Eva KW, Colliver JA. Theory development and application in medical education. *Teach Learn Med*. 2013;25(s1):s75-s80. DOI: 10.1080/10401334.2013.842907
14. Chi MT, Glaser R, Rees E. Expertise in problem solving. In: Sternberg RJ (Hrsg). *Advances in the psychology of human intelligence*. Hillsdale NJ: Lawrence Erlbaum Associates; 1982. S.7-76.
15. Eva KW, Neville AJ, Norman GR. Exploring the etiology of content specificity: Factors influencing analogic transfer and problem solving. *Acad Med*. 1998;73(10):s1-5. DOI: 10.1097/00001888-199810000-00028
16. Regehr G, Norman GR. Issues in cognitive psychology: Implications for professional education. *Acad Med*. 1996;71(9):988-1001. DOI: 10.1097/00001888-199609000-00015
17. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Med Educ*. 2007;41(12):1124-1130. DOI: 10.1111/j.1365-2923.2007.02946.x
18. Ericsson KA, Charness N. Expert performance. *Am Psychol*. 1994;49(8):725-747. DOI: 10.1037/0003-066X.49.8.725
19. Roediger HL, Karpicke JD. Test-Enhanced Learning. *Psychol Sci*. 2006;17(3):249-255. DOI: 10.1111/j.1467-9280.2006.01693.x
20. Van der Vleuten CP, Verwijnen GM, Wijnen WH. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Med Teach*. 1996;18(2):103-110. DOI: 10.3109/01421599609034142
21. Schuwirth LW, Van der Vleuten CP. The use of progress testing. *Perspect Med Educ*. 2012;1(1):24-30. DOI: 10.1007/s40037-012-0007-2
22. Hommes J, Van den Bossche P, de Grave W, Bos G, Schuwirth L, Scherpbier A. Understanding the effects of time on collaborative learning processes in problem based learning: a mixed methods study. *Adv Health Sci Educ Theory Pract*. 2014;19(4):541-563. DOI: 10.1007/s10459-013-9487-z
23. Hommes J, Rienties B, De Grave W, Bos G, Schuwirth L, Scherpbier A. Visualising the invisible: a network approach to reveal the informal social side of student learning. *Adv Health Sci Educ Theory Pract*. 2012;17(5):743-757. DOI: 10.1007/s10459-012-9349-0
24. Mezirow J. *Transformative Learning: Theory to Practice*. New Dir Adult Cont Educ. 2002;74:5-12.
25. Davis D, Thomson M, Oxman A, Haynes R. Changing Physician Performance: A Systematic Review of the Effect of Continuing Medical Education Strategies. *JAMA*. 1995;274(9):700-705. DOI: 10.1001/jama.1995.03530090032018
26. Pintrich P. A Conceptual Framework for Assessing Motivation and Self-Regulated Learning in College Students. *Educ Psychol Rev*. 2004;16(4):385-407. DOI: 10.1007/s10648-004-0006-x

**Corresponding author:**

Prof. Dr. Lambert Schuwirth  
 Flinders Universität, School of Medicine, Sturt Road,  
 Bedford Park, 5042 GPO Box 2100, | Adelaide SA 5001,  
 Australia  
[lambert.schuwirth@flinders.edu.au](mailto:lambert.schuwirth@flinders.edu.au)

**Please cite as**

Schuwirth L, Valentine N, Dilena P. An application of programmatic assessment for learning (PAL) system for general practice training. *GMS J Med Educ*. 2017;34(5):Doc56.  
 DOI: 10.3205/zma001133, URN: urn:nbn:de:0183-zma0011337

**This article is freely available from**

<http://www.egms.de/en/journals/zma/2017-34/zma001133.shtml>

**Received:** 2016-10-18

**Revised:** 2017-03-08

**Accepted:** 2017-05-08

**Published:** 2017-11-15

**Copyright**

©2017 Schuwirth et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

# Die Anwendung eines programmatischen Assessmentsystems (PAL) bei der (Facharzt)-Ausbildung von Allgemeinmedizinern

## Zusammenfassung

**Ziel:** Die programmatische Leistungsmessung bei Lernvorgängen (programmatic assessment for learning=PAL) wird als Konzept immer beliebter, jedoch ist dessen Umsetzung nicht unproblematisch. In dieser Arbeit beschreiben wir die Design Prinzipien, die einem PAL-Programm im Kontext der Facharzt-Ausbildung von Allgemeinmedizinern zugrunde liegen.

**Design Prinzipien:** Das PAL-Programm wurde entworfen, um die Aussagekraft von Informationen zur Beurteilung der Leistungen eines Weiterzubildenden zu optimieren und diese dazu zu bewegen, diese Informationen zu nutzen, das eigene Lernen selbst zu steuern. Die Hauptprinzipien des Programms sind kognitivistischer und transformativer Art. Die wesentlichen kognitiven Prinzipien, die wir dabei verwendeten, förderten das Verständnis tieferer Strukturen und stimulierten den (Wissens-)Transfer, indem die Weiterzubildenden kontinuierlich dazu angehalten wurden Praxiserfahrungen mit Hintergrundwissen zu verbinden. Ericssons Prinzip der bewussten Praxis wurde im Hinblick auf das Feedback zu Leistungserhebungen integriert und mit Pintrichs Modell der Selbststeuerung kombiniert. Mezirows transformatives Lernen und die Einblicke in die Sozialnetzwerktheorie zu kollaborativem Lernen wurden dazu verwendet, um die Weiterzubildenden in ihrer Entwicklung hin zu professionellen Fachärzten der Allgemeinmedizin zu unterstützen. Zu guter Letzt wurde noch das Prinzip des Test-basierten Lernens verbessert.

**Epilog:** Wir präsentieren dieses Beispiel, um die Design Entscheidungen, die unserem Programm zugrunde lagen, zu erläutern; jedoch möchten wir unser Programm nicht als (Patent-)Lösung für jede beliebige Situation verstanden wissen.

**Schlüsselwörter:** Examinierung, Hasuartausbildung, Programmatisches Assessment

## Hintergrund

Die programmatische Leistungsmessung bei Lernvorgängen (programmatic assessment for learning=PAL) gewinnt weltweit rapide an Popularität [1], [2], [3]. Das ist ziemlich überraschend, da sich das Konzept fundamental von dem unterscheidet, was in der Vergangenheit im Assessment-Bereich üblich war. Traditionell konzentrierte sich Assessment fast gänzlich auf die Feststellung, ob ein Student genügend gelernt hatte, um zu verhindern, dass noch-nicht-kompetente Studenten den nächsten Ausbildungsschritt erreichen. Wenn es einen Einfluss von Leistungsmessung auf studentisches Lernen gab, dann hauptsächlich vom behavioristischen Blickwinkel aus; das heißt, das Bestehen einer Prüfung war das sprichwörtliche „Zuckerbrot“ und das Nichtbestehen war die

„Peitsche“. Die Art und Weise, wie Prüfungen und deren Bewertung das Lernen von Studenten beeinflusst, ist jedoch weitaus komplexer. Cilliers et al. zum Beispiel zeigten die zahlreichen Wechselwirkungen zwischen den Ursprüngen der Einflüsse auf das Lernen, den Mechanismen, durch die das Lernen beeinflusst wird, und den möglichen Konsequenzen [4], [5]. Dieses Verständnis der Beziehung zwischen Leistungsmessung und Lernverhalten von Studenten ist wichtig bei der Anwendung von Assessment, insbesondere, um Studenten beim Lernen auf sinnvollere Weise anzuleiten. Das ist der Hauptzweck des Assessments für das Lernen [6], [7].

Um durch Leistungsmessung sinnvolles Lernen von Studenten zu lenken, müssen Studierende dazu befähigt werden wichtige Informationen, welche sie über ihre eigenen (Lern-)Leistungen erhalten, für ihr zukünftiges Lernen zu nutzen. Hier kommt der programmatische Aspekt des PAL ins Spiel [8]. Lediglich eine Punktzahl und ein Ergeb-

Lambert Schuwirth<sup>1,2,3,4</sup>  
Nyoli Valentine<sup>5</sup>  
Paul Dilena<sup>5</sup>

1 Flinders Universität, Adelaide, Australien

2 Universität Maastricht, Maastricht, Niederlande

3 Chang Gung Universität, Taiwan

4 Uniformed Services University, USA

5 Sturt Fleurieu GP, GPEx, Australien

nis über Bestanden/Nicht-Bestanden mitzuteilen, bietet den Studierenden nicht ausreichende Informationen über ihre Stärken und Schwächen. D.h. es hilft ihnen nicht dabei, spezifischere und konkretere Lernpläne zu formulieren. In vielen traditionellen Auswertungsprogrammen basiert das Kombinieren von Informationen auf dem Format des Assessments; z. B. wird in einem OSCE die Leistung bei einer abdominalen Untersuchungsstation mit der Leistung bei einer Knieuntersuchung verbunden. Bei PAL werden Ergebnisse von verschiedenen Auswertungsmethoden miteinander kombiniert, um sie aussagekräftiger zu machen. Dieses Prinzip wird vielleicht am besten an einem klinischen Beispiel illustriert [9]. Wenn man die Beschwerden eines Patienten wie Müdigkeit, Durstgefühl und häufiges Wasserlassen mit körperlichen Untersuchungsergebnissen, wie z. B. schlechte Wundheilung und fehlender arterieller Puls, sowie mit dem Blutglukosewert von 32 mmol pro Liter verbindet, gelangt man leicht zur Diagnose „Diabetes Mellitus“. Wir würden nicht in Erwägung ziehen, einem Patienten mitzuteilen, dass sein Glukosewert viel zu hoch sei, aber glücklicherweise sein Kaliumwert viel zu niedrig, so dass es ihm im Durchschnitt gut gehe – was die klinische Entsprechung für die Kombination der Leistung bei einer abdominalen Untersuchung mit der Leistung bei einer Knieuntersuchung wäre. Bei PAL geht es uns vielmehr darum, Ergebnisse über verschiedene Leistungsmessungsmethoden hinweg zu kombinieren, um so auf gleiche Weise zur Diagnose „Dyskompetenz“ zu gelangen. Selbstverständlich existiert die Krankheit „Dyskompetenz“ nicht; wir verwenden diesen Begriff lediglich, um das Prinzip zu illustrieren/veranschaulichen.

Ein zusätzliches Merkmal von PAL ist das Konzept der Verhältnismäßigkeit. Das bedeutet, dass die Art und Weise, wie Entscheidungen bezüglich der Leistung von Studierenden getroffen werden, proportional zur Plausibilität der zugrundeliegenden Informationen sein muss. Daher können einzelne Beobachtungen oder Einzelbewertungen für ein Feedback verwendet werden, aber nicht für weitreichende Entscheidungen. Bei PAL werden alle Einzelbeobachtungen oder Einzelbewertungen über einen Zeitraum hinweg gesammelt und zusammengetragen, bis ausreichend Informationen zur Verfügung stehen, um eine weitreichende Entscheidung bezüglich der Leistungen der Studierenden zu treffen [10]. Dies wiederum ist der täglichen klinischen Praxis sehr ähnlich; wir haben keine Bedenken, auf der Grundlage weniger Informationen eine einfache Diagnose zu stellen – z. B. in Bezug auf eine Infektion der oberen Atemwege – aber bei einer schwerwiegenden Diagnose – wie etwa eines bösartigen Tumors – ist es notwendig, sich auf verschiedene diagnostische Informationen stützen (Laborwerte, Bildgebung, Pathologie etc.).

Als logische Schlussfolgerung ist PAL damit ein longitudinaler Ansatz der Leistungsmessung, bei dem die Ergebnisse vieler formeller und informeller Leistungserhebungen kontinuierlich gesammelt werden, beispielsweise in einem Portfolio. Üblicherweise treffen sich der Lernende und ein Facharzt – oft als Mentor oder Coach bezeichnet

– in regelmäßigen Abständen, um die Fortschritte des Lernenden und dessen konkrete Lernziele zu besprechen. Und am Ende der Ausbildungsphase werden alle Ergebnisse evaluiert, um über das weitere fachbezogene Vorkommen der Weiterzubildenden zu entscheiden. Eine wichtige Voraussetzung dafür ist, dass Lernende auf Basis aller zur Verfügung stehenden Informationen bezüglich ihrer Lernleistungen eine Analyse durchführen und konkrete Lernziele formulieren, bevor sie sich mit ihrem Mentor oder Coach treffen. Auf diese Weise ist es den Mentoren möglich, bereits während der Ausbildungsphase eine Prognose hinsichtlich des möglichen Ausbildungsergebnisses abzugeben. Das Konzept von PAL wurde in verschiedenen Veröffentlichungen beschrieben [1], [7], [8], [10].

Wenn man das Konzept des PAL erläutert – insbesondere Gesundheitsdienstleistern gegenüber – ist es in der Regel so, dass sich das Konzept intuitiv richtig anfühlt und die Verantwortlichen zustimmen, jedoch gestaltet sich dessen Umsetzung als äußerst schwierig. Dafür gibt es mehrere Gründe. Der erste und wohl wichtigste Grund ist die Tatsache, dass es eine fundamental veränderte Sichtweise auf die Rolle von Leistungsmessung darstellt. Solche fundamentalen Veränderungen brauchen in jeder Fachrichtung Zeit, um ihren Weg von der Theorie in die Praxis zu finden. So benötigten konzeptuell andere Denkansätze in der Medizindidaktik, wie das Problem-basierte Lernen, mehrere Jahre oder gar Jahrzehnte, bevor sie weithin akzeptiert wurden. Ein zweiter Grund betrifft die logistischen Veränderungen, die für eine Umsetzung des PAL nötig sind. Denn das Programm ist in seinem Aufbau explizit, weshalb der damit zusammenhängende Aufwand in Bezug auf Zeit und Kosten offensichtlich wird, wohingegen bei vielen traditionellen Assessmentprogrammen die Kosten im Allgemeinen eher verdeckt sind. Dies gestaltet einen Kostenvergleich ziemlich schwierig und führt leicht zu einer negativen Wahrnehmung von PAL. Der dritte Grund liegt aller Wahrscheinlichkeit nach darin begründet, was Vosniadou als naive Rahmenvorgaben bzw. naive Theorien bezeichnet [11]. Aufgrund unserer Lebenserfahrungen entwickeln wir eine eigene Sichtweise darauf, wie die Welt funktioniert, und es ist sehr schwierig, diese Überzeugungen zu ändern. Sie können durch formellere Theorien ergänzt werden, aber sie verschwinden nie wirklich ganz. Dasselbe findet sich zum Thema Ausbildung. Unsere Meinung darüber, was Ausbildung ist und wie sie gestaltet werden sollte, wurde durch unsere eigenen langjährigen Erfahrungen als Schüler und Studierende geprägt und obwohl diese Ansichten durch didaktische Theorien und Ansätze modifiziert werden können – z. B. Personalentwicklung oder Lehrerausbildung – werden sie nie völlig verschwinden. Wenn daher ein Versuch unternommen wird, PAL umzusetzen, beeinflussen die naiven Überzeugungen weiterhin dessen verschiedene Design-Entscheidungen.

Einige der eher intuitiven Herangehensweisen im Umgang mit derartigen Grundüberzeugungen und deren Wandel betreffen eine sorgfältige Identifizierung von Akteuren und deren jeweiliger Funktion sowie die vorsichtige

Kommunikation mit ihnen. Diese Kommunikation muss offen und kontinuierlich verlaufen, aber auch flexibel, damit sie auf die verschiedenen Akteure und deren unterschiedliche Argumente und Erklärungen eingehen kann; von Evidenz aus der Forschung bis hin zu rhetorischer Überzeugung. Ein weiterer Faktor, der bei diesem Prozess helfen kann, ist die Beschreibung der Art und Weise, wie PAL andernorts umgesetzt wurde, um zu zeigen, wie das Konzept erfolgreich in die Praxis überführt werden kann. Sinn und Zweck dieses Artikels ist es, einen Nachweis über die Umsetzung des Konzepts aufzuzeigen. In der Konsenserklärung und den Empfehlungen der Konferenz von Ottawa im Jahr 2010 [12] wird die ideographische Beschreibung bzw. der Fallbericht als eine wichtige Art der Forschung gesehen, sofern die beschriebene Praxis mit den zugrunde liegenden theoretischen Konzepten in Beziehung gesetzt wird, damit die Leser die Gestaltungsentscheidungen verstehen und ihrem eigenen Kontext anpassen können. In der ärztlichen Ausbildung wird diese „Anpassungsfähigkeit“ als hilfreicher angesehen, als die bloße Replizierbarkeit von Ergebnissen [13]. Aus diesem Grund stellt dieser Aufsatz eine Fallstudie dar.

## Kontext

Das GP365 ist ein Ausbildungsprogramm für angehende Fachärzte im Bereich der Allgemeinmedizin in Süd- und West-Australien, das von Sturt Fleurieu Education and Training in Zusammenarbeit mit dem Prideaux Centre for Research in Health Professions Education an der Flinders Universität entwickelt wurde. Es handelt sich um ein einjähriges Curriculum im Kontext eines Ausbildungsprogramms mit der Gesamtdauer von drei bis vier Jahren, welches allen Allgemeinmedizinern in der klinischen Facharztausbildung in Süd und West-Australien offensteht. Während dieses einen Jahres folgen die Weiterzubildenden dem GP365-Programm, das sie bei der Verknüpfung von Hintergrundwissen, Fertigkeiten und Verständnis mit ihren praktischen Erfahrungen unterstützt. In diesem Jahr stellt GP365 den Weiterzubildenden Hintergrundstudienmaterial, verschiedene Aufgaben zur Bearbeitung, einen Mentor, einen Medical Educator sowie eine Peer Group zur Verfügung, mit denen die Weiterzubildenden zusammenarbeiten. Die Weiterzubildenden erhalten dabei regelmäßig Feedback von ihren Mentoren, Medical Educators und Peers. Zusätzlich erhalten sie Übungsaufgaben, um relevantes Wissen und dessen Anwendung zu prüfen. Während ihrer Ausbildung erstellen die Weiterzubildenden ein Portfolio, das am Ende Nachweise über alle Arten von Feedback und Bewertungen enthält, wie: direkt beobachtete Patientenberatungen, begutachtete Videoaufzeichnungen von Beratungsgesprächen, kritische Fallanalyse-Berichte, klinische Audits, Professionalität, Aktivitäten in der Peer Group, Multi-Source-Feedback, Prüfungen während und am Ende des Ausbildungsabschnitts, Ergebnisse der Übungsaufgaben zusammen mit eigenen Analysen sowie schriftliches Feedback von Medical Educators.

Insgesamt betrachtet mag dieses Programm weder innovativ noch besonders anders im Vergleich zu vielen postgraduellen Ausbildungskontexten erscheinen, aber es bestehen Unterschiede, die wir im Nachfolgenden erläutern.

## Design Prinzipien

Das wichtigste Design Prinzip das GP365 zugrunde liegt ist das Problem des (Wissens-)Transfers und des Verstehens der sogenannten „tiefen Struktur“ [14]. Weiterzubildende erleben während ihrer Facharztausbildung eine Vielzahl einzelner Patientenfälle, um jedoch ein Experte zu werden, ist es wichtig, Ähnlichkeiten und Unterschiede zwischen diesen Fällen zu verstehen, also (Wissens-)Transfer zu vollbringen [15]. In der Literatur wird die Bedeutung von Dekontextualisierung, des Verständnisses erster Prinzipien und von Rekontextualisierung (Anwendung dieser Prinzipien auf einen anderen Fall) für die Entwicklung von (Wissens-)Transfer und Fachwissen beschrieben [15], [16]. Daher richtet sich beim GP365 die Leistungsmessung danach, Weiterzubildende dabei zu unterstützen, diesen individuellen Erfahrungen eine tiefere Bedeutung zukommen zu lassen, indem beispielsweise von ihnen verlangt wird, medizinisches Grundwissen und klinisches Hintergrundwissen mit ihren individuellen Patientenerfahrungen in Beziehung zu setzen. Das ist typischerweise das, worauf sich kritische Fallanalyse-Berichte konzentrieren. Der Weiterzubildende wählt einen Patienten für seinen kritischen Fallanalyse-Bericht aus, muss aber in der Lage sein, zu erklären, warum er gerade diesen Patienten für sein eigenes Lernen als ganz besonders relevant ansieht. Er definiert zudem seine eigenen konkreten Lernziele und wertet dann die erforderlichen Hintergrundinformationen aus, um zu einem umfassenden Verständnis des klinischen Falls und dessen Behandlung zu gelangen. Als „Nachweis“ dieses Lernvorgangs legt der Weiterzubildende drei fallbasierte Multiple-Choice-Fragen vor, die mit Verweisen auf die Fachliteratur gestützt werden. Er bzw. sie reicht danach den Fallbericht und die Fragen ein. Der Medical Educator liest den klinischen Fallbericht und gibt ein umfangreiches Feedback, das der Weiterzubildende in eine überarbeitete Version einarbeiten muss. Diese Vorgehensweise erfordert vom Betreffenden, in optimaler Weise das in der klinischen Praxis Erlebte in einen sinnvollen Bezug zur Theorie zu setzen. Dieser Leistungsnachweis wird überprüft und verpflichtender Bestandteil des Portfolios – so wie alle anderen Arten der Leistungsmessung – und fließt in die endgültige Entscheidung mit ein.

Das zweite Design Prinzip ist die Ansammlung von Informationen über verschiedene Assessment-Teile hinweg. Ein Beispiel für eine solche Verbindung beginnt mit kritischen Fallanalyse-Berichten. Wie oben beschrieben, erhält der Weiterzubildende ein Feedback zu all seinen Berichten, und muss die Berichte entsprechend überarbeiten. Das ist eine Anwendung von Ericssons Prinzip der bewussten Praxis [17], [18]. Die Multiple-Choice-Fragen,



die jeder Weiterzubildende ausarbeitet, werden in einer Fragen-Datenbank gesammelt. Auf Grundlage dieser Datenbank werden regelmäßig Tests zum Lernfortschritt zusammengestellt und den Weiterzubildenden vorgelegt, die diese dann innerhalb eines festgelegten Zeitfensters unter Verwendung des digitalen Lernsystems bearbeiten müssen. Die Testfragen werden dann freigegeben, und die Weiterzubildenden müssen mindestens drei Fragestellungen kritisch erörtern, vorzugsweise solche Fragen, die ihnen am umstrittensten erscheinen. Bei der kritischen Erörterung dieser Fragen, müssen sie mit relevanter wissenschaftlicher Literatur ihren Standpunkt untermauern. Die Idee dahinter ist, den Einfluss des Test-basierten Lernens zu optimieren, indem die Weiterzubildenden dazu angehalten werden, die Fragestellungen und ihre eigenen Stellungnahmen dazu kritisch zu hinterfragen [19]. Diese Methode ist in verschiedenen Situationen zur Überprüfung des Lernfortschritts angewendet worden [20], [21]. Diese kritischen Erörterungen werden bei Treffen der Peer Group besprochen. Nach dem Austausch dieser kritischen Standpunkte zwischen den Gruppenmitgliedern, wird von der Gruppe erwartet, dass sie einen Konsens zu den noch strittigen Fragen erarbeitet und diese abschließend zusammenfasst. Nur so erhält der Weiterzubildende seine Ergebnisse für die Tests, die er dann auswerten und in seinem Portfolio verwenden kann. Das Prinzip bei diesen Gruppentreffen ist es, die Entwicklung von informellen Peer-Netzwerken zu fördern. Weiterzubildende arbeiten vielleicht in abgelegenen Gegenden, wo viele evtl. nur einen eingeschränkten Kollegenkreis oder gar keine Kollegen in ihrem Alter und mit ihrem Erfahrungsstand haben, an die sie sich wenden könnten. Die Literatur zeigt, dass das Vorhandensein informeller Netzwerke wichtig für den Erwerb von Informationen und für das Lernen selber ist [22], [23]. Ein Prinzip der transformativen Lerntheorie, das ebenfalls miteinbezogen wurde, konzentriert sich darauf, Weiterzubildenden bewusst zu machen, dass nicht alles, was in der Fachliteratur geschrieben steht ohne Zweifel richtig ist, und dass Toleranz für eine gewisse Unsicherheit Teil der Praxis ist [24].

Dies dient der Veranschaulichung, wie das Assessment-Programm Weiterzubildende dazu bewegt, die Informationen zur Beurteilung ihrer Leistungen über verschiedene Wege und auf sinnvollere Art und Weise zu integrieren, um damit die „konstruktivistischen“ Komponenten der zielgerichteten Auswertung von Lernvorgängen zu optimieren.

Ein drittes Prinzip ist der Zuwachs an Eigenverantwortung für das (selbstgesteuerte) Lernen. Im Allgemeinen wird von Studierenden nach dem Universitätsabschluss erwartet, dass sie in der Lage sind, ihr eigenes Lernen und die Beurteilung ihrer Lernleistungen selbst zu kontrollieren. Leider ist das nicht immer der Fall. Eines der Probleme bei CME ist beispielsweise, die Neigung von Menschen, Kurse in solchen Bereichen zu belegen, in denen sie bereits gut sind [25]. Die Beurteilung und Auswertung für das Lernen soll Lernende eigentlich dazu befähigen, ihre Stärken und Schwächen zu analysieren, um diese dann

in spezifische Lernziele zu übertragen und diese Lernstrategien tatsächlich zu realisieren. Für die meisten entsteht selbst gesteuertes Lernen nicht von selbst, und es bedarf eines Entwicklungsprozesses und der gezielten Anleitung. Paul Pintrich's Modell ist hilfreich, da es Prozesse wie „Vorüberlegung, Planung und Aktivierung“, „Überwachung“, „Kontrolle“ und „Reaktion und Reflexion“ voneinander trennt; jede(r) davon verlangt vom Lernenden die aktive Auseinandersetzung mit seiner Wahrnehmung, seiner Motivation, seines Verhaltens und des Kontexts [26]. Selbstverständlich sind die Elemente „Vorüberlegung, Planung und Aktivierung“, „Überwachung“, „Kontrolle“ und „Reaktion und Reflexion“ in das oben dargestellte System der Leistungsmessung integriert, indem die angehenden Fachärzte an regelmäßigen Treffen teilnehmen und kontinuierlich dazu angehalten werden, ihre eigenen Fortschritte, Stärken und Schwächen zu analysieren, und daraus konkret realisierbare Lernziele zu formulieren. Weiterzubildende, die diese selbstgesteuerten Schritte nicht oder nicht in ausreichendem Umfang vornehmen, müssen diese Prozesse nochmals durchlaufen und es wird ihnen möglicherweise nicht erlaubt, die nächste Ausbildungsphase zu beginnen. Indem den Weiterzubildenden ein Feedback, eine Peer Group, ein Mentor und ein engagierter Medical Educator zur Verfügung gestellt werden, unterstützt das Programm die Motivation und das Verhalten der angehenden Fachärzte. Zudem unterstützt es ihre Wahrnehmung, indem es umfangreiche Information während der Phase der Leistungsmessungen bereit stellt. Ihr Verhältnis zu ihrem Mentor, ihrem Medical Educator und ihrer Peer Group bietet ihnen dabei Unterstützung zu lernen, wie sie sich in ihrem aktuellen und zukünftigen komplexen beruflichen Kontext bewegen können. Weiterzubildende, die nicht die gewünschten Fortschritte erzielen, müssen sich einer Fördermaßnahme unterziehen. Aber sie selbst müssen die Initiative ergreifen und ihren eigenen Plan ausarbeiten, wie diese Fördermaßnahmen gestaltet sein sollen; selbstverständlich mit Unterstützung und Erlaubnis ihrer Mentoren und Medical Educators. Dabei obliegt ihnen die alleinige Verantwortung für ihre eigenen Lernstrategien, so, wie es von ihnen auch nach dem Studienabschluss verlangt wird.

Ein viertes Prinzip ist die longitudinale Qualität des Programms. Sämtliche Informationen über die Leistungen des Weiterzubildenden werden in einem Portfolio gesammelt, das regelmäßig mit den Mentoren und/oder Medical Educators besprochen wird. Am Anfang, wenn die Informationen im Portfolio noch „dünn“ sind, wird meist ein eher formatives Feedback gegeben; erst wenn die Informationen ergiebiger sind, werden ernsthafte Vorschläge für Interventionen und konkrete Hilfestellung gemacht, und die endgültige Entscheidung, ob ein Weiterzubildender bereit ist, zur nächsten Phase fortzuschreiten, wird immer auf Grundlage aller aussagekräftigen Informationen hinsichtlich seiner Leistungen getroffen.

Ein abschließendes Design Prinzip ist die Anpassungsfähigkeit des Programms an den jeweiligen lokalen Kontext. Es ist unwahrscheinlich, dass jeder Ansatz in der Medizi-

nerausbildung, der in einem bestimmten Land oder Kontext gut funktioniert, einfach auf ein anderes Land übertragen werden kann. Um Bildung erfolgreich zu gestalten, muss sie mit den Erwartungen und den kulturellen Gegebenheiten ihrer Umgebung verknüpft werden. Beim GP365 war es daher wichtig den Umstand zu berücksichtigen, dass Australien ein sehr großes Land mit vielen abgelegenen Gegenden ist. Allgemeinärzte spielen im Kontext der australischen Gesundheitsvorsorge eine wichtige Rolle, und fast immer sind sie die erste Anlaufstelle (bei gesundheitlichen Beschwerden); zudem sind sie oftmals auch die einzige Anlaufstelle. Bei den Ausbildungsprogrammen für Allgemeinärzte wird somit besonderer Wert darauf gelegt Allgemeinärzte auszubilden, die bestens darauf vorbereitet sind, selbständig und sicher zu arbeiten. Deshalb sind Elemente, wie selbstgesteuertes Lernen, Verantwortlichkeit, der Umgang mit Unsicherheiten und Ambiguitätstoleranz hervorsteckende Kennzeichen des Programms, wie zum Beispiel die Treffen der Peer Groups – die sogenannten Mini-Selbsthilfegruppen – die laufenden Feedbacks zur Professionalität sowie die regelmäßigen Treffen mit erfahrenen Mentoren und Medical Educators. Das alles ist in verschiedensten Aspekte mit dem Programm verwoben, um sicherzustellen, dass das Programm den Anforderungen der Colleges hinsichtlich der Vorgaben „sicher und selbständig arbeitenden Allgemeinmediziner“ und „Colleg-Examen-Bereitschaft“ entspricht.

## Epilog

Das Programm läuft seit nunmehr zwei Jahren und wird bald umfassende Resultate zu seinem Erfolg hinsichtlich der Qualität der Absolventen aufzeigen können. Frühe Ergebnisse deuten an, dass die die Quote der Absolventen mit erfolgreichem Abschluss bei den Fellowship-Prüfungen deutlich über dem landesweiten Durchschnitt liegt, aber bevor eine klare kausale Beziehung zum PAL-Ansatz hergestellt werden kann, sind viele weitere Daten notwendig. Obwohl es anekdotische Anhaltspunkte gibt, dass die programmatische Leistungsmessung bei Lernvorgängen effektiv ist, ist der gesamte Ansatz noch zu jung, um hinreichend zuverlässige Evidenz zu erbringen. Ein endgültiges Urteil steht also noch aus. Darüber hinaus sind wir der Ansicht, dass ein Erfolg in unserem Kontext noch nicht bedeutet, dass genau das gleiche Programm in einem anderen Kontext gleichermaßen erfolgreich wäre. Deshalb haben wir uns darauf konzentriert, Beispiele zu zeigen, wie das PAL-Konzept bei Bewertung von Leistungen die Design Entscheidungen rund um das GP365 Assessment-Programm beeinflusst hat. Wir hoffen, dass unsere Erläuterungen der zugrundeliegenden Prinzipien sowie die Beschreibung, wie diese unsere Design Entscheidungen beeinflusst haben, für all jene hilfreich sind, die darüber nachdenken, PAL im Rahmen Leistungsmessung zu verwenden.

## Interessenkonflikt

Die Autoren erklären, dass sie keine Interessenkonflikte im Zusammenhang mit diesem Artikel haben.

## Literatur

1. Van der Vleuten CP, Schuwirth LW. Assessing professional competence: from methods to programmes. *Med Educ*. 2005;39(3):309-317. DOI: 10.1111/j.1365-2929.2005.02094.x
2. Fishlede A, Henson L, Hull A. Cleveland Clinic Lerner College of Medicine: An innovative approach to medical education and the training of physician investigators. *Acad Med*. 2007;82(4):390-396. DOI: 10.1097/ACM.0b013e318033364e
3. Schuwirth LW, Ward H, Heeneman S. Assessment for Learning. In: Higgs J, Baldry-Currens J, Jensen G, Letts W, Sheehan D (Hrsg). *Realising Exemplary Practice-based Education*. Rotterdam: Sense Publishers; 2013. S.143-150. DOI: 10.1007/978-94-6209-188-7\_15
4. Cilliers FJ, Schuwirth LW, Herman N, Adendorff HJ, Van der Vleuten CP. A model of the pre-assessment learning effects of summative assessment in medical education. *Adv Health Sci Educ Theory Pract*. 2012;17(1):39-53. DOI: 10.1007/s10459-011-9292-5
5. Cilliers FJ, Schuwirth LW, Adendorff HJ, Herman N, Van der Vleuten CP. The mechanisms of impact of summative assessment on medical students' learning. *Adv Health Sci Educ Theory Pract*. 2010;15(5):695-715. DOI: 10.1007/s10459-010-9232-9
6. Shepard L. The role of assessment in a learning culture. *Educ Res*. 2009;29(7):4-14. DOI: 10.3102/0013189X029007004
7. Schuwirth L. From assessment of learning to assessment for learning. Huntington WV: IAMSE; 2107. S.170. Zugänglich unter/available from: <http://www.iamse.org>
8. Schuwirth LW, Van der Vleuten CP. Programmatic assessment: from assessment of learning to assessment for learning. *Med Teach*. 2011;33(6):478-485. DOI: 10.3109/0142159X.2011.565828
9. Schuwirth LW, Van der Vleuten CP, Durning SJ. What programmatic assessment for learning in medical education can learn from healthcare. *Perspect Med Educ*. 2017;6(4):211-215. DOI: 10.1007/s40037-017-0345-1
10. Van der Vleuten CP, Schuwirth LW, Scheele F, Driessen EW, Hodges B. The assessment of professional competence: building blocks for theory development. *Best Pract Res Clin Obstet Gynaecol*. 2010;24(6):703-719. DOI: 10.1016/j.bpobgyn.2010.04.001
11. Vosniadou S. Capturing and modeling the process of conceptual change. *Learn Instruct*. 1994;4:45-69. DOI: 10.1016/0959-4752(94)90018-3
12. Schuwirth LW, Colliver J, Gruppen L, Mennin S, Kreiter C, Mennin S, Onishi H, Pangaro L, Ringsted C, Swanson D, Van der Vleuten C, Wagner-Menghin M. Research in assessment: Consensus statement and recommendation from the Ottawa 2010 conference. *Med Teach*. 2011;33(3):224-233. DOI: 10.3109/0142159X.2011.551558
13. Cianciolo AT, Eva KW, Colliver JA. Theory development and application in medical education. *Teach Learn Med*. 2013;25(s1):s75-s80. DOI: 10.1080/10401334.2013.842907
14. Chi MT, Glaser R, Rees E. Expertise in problem solving. In: Sternberg RJ (Hrsg). *Advances in the psychology of human intelligence*. Hillsdale NJ: Lawrence Erlbaum Associates; 1982. S.7-76.

15. Eva KW, Neville AJ, Norman GR. Exploring the etiology of content specificity: Factors influencing analogic transfer and problem solving. *Acad Med.* 1998;73(10):s1-5. DOI: 10.1097/00001888-199810000-00028
16. Regehr G, Norman GR. Issues in cognitive psychology: Implications for professional education. *Acad Med.* 1996;71(9):988-1001. DOI: 10.1097/00001888-199609000-00015
17. Ericsson KA. An expert-performance perspective of research on medical expertise: the study of clinical performance. *Med Educ.* 2007;41(12):1124-1130. DOI: 10.1111/j.1365-2923.2007.02946.x
18. Ericsson KA, Charness N. Expert performance. *Am Psychol.* 1994;49(8):725-747. DOI: 10.1037/0003-066X.49.8.725
19. Roediger HL, Karpicke JD. Test-Enhanced Learning. *Psychol Sci.* 2006;17(3):249-255. DOI: 10.1111/j.1467-9280.2006.01693.x
20. Van der Vleuten CP, Verwijnen GM, Wijnen WH. Fifteen years of experience with progress testing in a problem-based learning curriculum. *Med Teach.* 1996;18(2):103-110. DOI: 10.3109/01421599609034142
21. Schuwirth LW, Van der Vleuten CP. The use of progress testing. *Perspect Med Educ.* 2012;1(1):24-30. DOI: 10.1007/s40037-012-0007-2
22. Hommes J, Van den Bossche P, de Grave W, Bos G, Schuwirth L, Scherpbier A. Understanding the effects of time on collaborative learning processes in problem based learning: a mixed methods study. *Adv Health Sci Educ Theory Pract.* 2014;19(4):541-563. DOI: 10.1007/s10459-013-9487-z
23. Hommes J, Rienties B, De Grave W, Bos G, Schuwirth L, Scherpbier A. Visualising the invisible: a network approach to reveal the informal social side of student learning. *Adv Health Sci Educ Theory Pract.* 2012;17(5):743-757. DOI: 10.1007/s10459-012-9349-0
24. Mezirow J. Transformative Learning: Theory to Practice. *New Dir Adult Cont Educ.* 2002;74:5-12.
25. Davis D, Thomson M, Oxman A, Haynes R. Changing Physician Performance: A Systematic Review of the Effect of Continuing Medical Education Strategies. *JAMA.* 1995;274(9):700-705. DOI: 10.1001/jama.1995.03530090032018
26. Pintrich P. A Conceptual Framework for Assessing Motivation and Self-Regulated Learning in College Students. *Educ Psychol Rev.* 2004;16(4):385-407. DOI: 10.1007/s10648-004-0006-x

**Korrespondenzadresse:**

Prof. Dr. Lambert Schuwirth  
 Flinders Universität, School of Medicine, Sturt Road,  
 Bedford Park, 5042 GPO Box 2100, | Adelaide SA 5001,  
 Australien  
 lambert.schuwirth@flinders.edu.au

**Bitte zitieren als**

Schuwirth L, Valentine N, Dilena P. An application of programmatic assessment for learning (PAL) system for general practice training. *GMS J Med Educ.* 2017;34(5):Doc56.  
 DOI: 10.3205/zma001133, URN: urn:nbn:de:0183-zma0011337

**Artikel online frei zugänglich unter**

<http://www.egms.de/en/journals/zma/2017-34/zma001133.shtml>

**Eingereicht:** 18.10.2016

**Überarbeitet:** 08.03.2017

**Angenommen:** 08.05.2017

**Veröffentlicht:** 15.11.2017

**Copyright**

©2017 Schuwirth et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.