

Educating the self-critical doctor

Using a portfolio
to stimulate and assess
medical students' reflection

Cover design: Zuiderlicht, Maastricht
ISBN: 978-90-78213-06-2

Educating the self-critical doctor

Using a portfolio
to stimulate and assess
medical students' reflection

Proefschrift

Ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus, Prof. mr. G.P.M.F. Mols,
volgens het besluit van het College van Decanen
in het openbaar te verdedigen op woensdag 25 juni 2008
om 16.00 uur

door

Erik W. Driessen
geboren op 28 oktober 1966, Wijchen

Promotoren

Prof.dr. J.D.H.M. Vermunt

Prof.dr. C.P.M. van der Vleuten

Co-promotor

Dr. J. van Tartwijk

Beoordelingscommissie

Prof.dr. A.J.J.A. Scherpbier (voorzitter)

Prof.dr. Th. J. ten Cate, Universiteit Utrecht

Prof.dr. E. Heineman

Dr. D.D. Mansvelder-Longayroux, Universiteit Leiden

Prof.dr. L.H.E.H. Snoeckx

For my parents, who would have been proud
and for my two sons, Sjako and Sijmen, of whom I am proud.

Contents

1	Introduction	9
2	Use of portfolios in early undergraduate medical training.....	25
	Published in: Medical Teacher 2003; 25: 18-23	
3	The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study	41
	Published in: Medical Education 2005; 39: 214-220	
4	Conditions for successful reflective use of portfolios in undergraduate medical education	55
	Published in: Medical Education 2005; 39: 1230-1235	
5	Validity of portfolio assessment: which qualities determine ratings?.....	67
	Published in: Medical Education 2006; 40: 862-866	
6	Web-based or paper-based portfolios: Is there a difference?.....	77
	Published in: Medical Education 2007; 41: 1067-1073	
7	Portfolios in medical education: why do they meet with mixed success? A systematic review.....	91
	Published in: Medical Education 2007; 41: 1224-1233	
8	The self-critical doctor: helping students become more reflective.....	123
	Published: British Medical Journal 2008; 336: 827-830	
9	Discussion.....	133
	Summary.....	145
	Samenvatting.....	151
	Curriculum Vitae	157

1 Introduction

Today's doctors find themselves confronted not only with patients who are increasingly knowledgeable and assertive but also with an explosion of medical knowledge and with pressure to apply new findings and evidence in day-to-day practice and collaborate with other health professionals in ever larger teams and communities.* To deal with these complexities, doctors need, in addition to the medical knowledge of their discipline, generic competencies to enhance effective communication, organisation, teamwork and professionalism. These generic competencies are sometimes labelled as doctors' 'soft skills' in contrast to 'hard clinical skills'. In recent years, learning, teaching and assessment of these generic competencies have gained unexpected urgency among politicians and the general public in the Netherlands¹ and elsewhere.² Headlines decrying incidents involving 'dysfunctional' doctors and hospital departments with dramatic impact on morbidity and mortality figures³ catapulted generic competencies to the forefront of attention as indispensable qualities for doctors. As a result, professional associations and governments began to voice increasingly urgent demands for including these generic competencies in education and assessment.^{1,4}

Generic communication, organisational and professional competencies can be acquired in a variety of ways, such as work-based task performance in different contexts, evaluation of work outcomes, identification of positive and negative results of work, and deliberate efforts to achieve learning objectives designed to address diagnosed weaknesses. Portfolios are considered to be the instrument of choice for supporting the development of generic competencies and for assessing whether they are being attained.^{5,6} Because the portfolio can provide visible and tangible evidence of how students are progressing, it is eminently suitable for monitoring, discussing and assessing student learning. In this thesis, we report the outcomes of a research project aimed at evaluating and improving a portfolio designed to support and assess medical students' progress in reflective skills. Before we

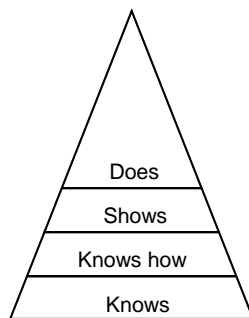
* Parts of this chapter were previously published in: van Tartwijk J, Driessen E, van der Vleuten C, Stokking K. Factors influencing the introduction of portfolios. *Quality in Higher Education*. (2007);13(1):69-79.

present the individual studies that together form this thesis, we will use this introduction to address 3 issues that are relevant to portfolios in medical education, namely the how portfolios developed, the diversity of portfolios, and the place of reflection in medical education.

The development of portfolios in medical education

In 1990, Miller presented a framework for clinical assessment, shaped like a pyramid, whose layers from bottom to top represent increasingly complex levels of mastery, with the lower levels providing the foundation for the higher levels.⁷

Figure 1.1 Miller's pyramid⁷



The bottom level is concerned with *knowledge*. This is the knowledge relating to the skills that students must master for their future professional practice. This knowledge is best assessed by written tests.

The next level represents application of the knowledge from level 1. Students should know *how* to apply their knowledge when performing skills. For instance, at this level, students are expected to know how to diagnose a patient and which aspects of a patient's presentation to attend to. The '*knows how*' level can also be assessed by written tests.

One level up, at level 3, the issue of interest is that students demonstrate their ability to use their knowledge to *take appropriate action in a simulated environment*. This level combines knowledge and action (cognition and behaviour). Not only should students know how to diagnose a patient, they should also be able to actually perform the appropriate actions, for example a physical examination in a simulated patient (shows how).

The top of the pyramid is concerned with independent *performance within the complex environment of day-to-day practice*. This requires integration of knowledge, skills, attitudes, and personal characteristics. Performance at the top of the pyramid is manifested when students are working independently in

professional practice. Typically, adequate performance at this level requires integrated performance of different roles. Not only the role of medical expert but also that of counsellor, participant in the doctor patient relationship, a leadership role in relation to nursing staff, et cetera. Good performance at the *Does* level implies competence. In higher education, the concept of *competence* is much used and much debated. Generally, it is defined as an integrated body of knowledge, skills, (professional) attitudes and traits enabling proficient performance in certain real life settings.⁸

In 1990, Miller observed that there were no instruments to evaluate performance consistent with the top of the pyramid. In the related but different field of teacher education and assessment, people were already struggling with the same problem in the late 1980s. Shulman describes the Teacher Assessment Project that was underway at Stanford University at that time.⁹ He recounts that it was considered undesirable to assess teacher competence solely on the basis of ratings in assessment centres. It was considered opportune to complement this type of assessment by judgements of whether teachers succeeded in making the most of their pupils' learning opportunities within their own complex working environment. Teachers may do so in different, teacher-specific ways, none of which should be seen as the single correct one. It was also recognised that there can be striking variations among teaching settings and that it does make a difference whether one teaches at a school in a deprived area of central Chicago with its myriad of social problems or at a high school in a middle class suburban area. As part of efforts to achieve fair judgement of teacher performance in a broad array of settings and situations, the *portfolio* concept was borrowed from the arts and architecture.

Originally, the term portfolio denoted a portable case for keeping, usually without folding, loose sheets of papers, drawings, photographs, et cetera.¹⁰ Artists, architects and designers have traditionally collected samples of their work in a portfolio, which they present to prospective clients to convince them of the quality of their work. In an essay published in 1990, Bird suggested that similar portfolios to those of artists and architects might make a suitable instrument for student assessment.¹¹ Portfolios could hold copies of materials, tests, photographs, observation reports, videotapes, handwritten notes, reports of evaluations, and so forth. Compared to assessment instruments like paper-and-pencil tests, assessment centre procedures and low-inference observation schedules tied to lists of students competencies, it was hoped that portfolios would offer an opportunity for assessors to see the

wide range of a student's work and incorporate in their judgements the limitations and opportunities of varying performance contexts. Bird's essay marked the start of a new trend in higher education: portfolios as a tool for assessing task performance in authentic situations.

Portfolio as a multipurpose instrument

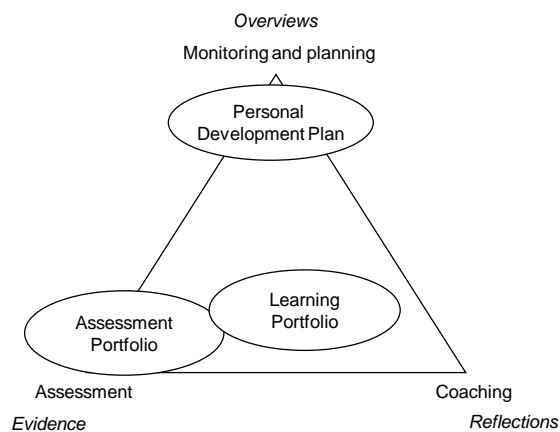
In the years following its debut as an assessment tool, the portfolio was advertised as an excellent instrument not only for authentic assessment but also for stimulating reflective thinking.⁵ Working on a portfolio can activate reflection, because collecting work samples, evaluations and other types of illustrative materials compels students to look back on what they have done and analyse what they have and have not accomplished. Another way in which portfolios can stimulate reflective thinking is through written reflections, whose inclusion is mandatory for many portfolios. Examples are: reflective journals or diaries,^{12,13} reflective essays,¹⁴ mission statements, self-evaluations, and descriptions of steps taken to achieve improvement.¹⁵ In many cases, portfolios are assembled over a longer period of time. That is why they can also be used as an instrument to support planning and monitoring in professional development. One way to do so is to include learning objectives in the portfolio and a document trail of related learning activities and accomplishments.^{13,16,17}

As a consequence, reflections and overviews of personal development have secured a prominent place in many portfolios. Portfolios that are primarily geared to assessment will remain organised around artefacts and other kinds of materials, which provide 'evidence' of competencies. In portfolios that are primarily used to monitor and plan students' development, overviews will take centre stage. Portfolios whose primary objective is to foster learning by stimulating students to reflect on and discuss their development will be organised around students' reflections.

Inevitably, these developments have widened the applicability of the label 'portfolio' to a broad range of instruments. Some portfolios might equally aptly be labelled 'Personal Development Plan' or 'Reflective Essay'. Because of the tremendous variety in portfolios, careful and critical appraisal of the strengths and weaknesses of different portfolio systems is advisable before deciding which one to implement in a particular setting. The question to be answered is whether a certain portfolio is fit for its intended purpose.¹⁸ Just like shoes, portfolios come in different shapes and sizes.¹⁹ And just as someone else's shoes are unlikely to fit comfortably, portfolios tailored to one particular school or university may not fit into the educational configuration(s) of other schools or universities. An ill-fitting portfolio will

inevitably be discarded sooner or later. To assist in determining whether a portfolio is appropriate for its intended purpose the triangle in figure 1.2 helps to clarify the nature of a portfolio. It does so by inviting positioning of a portfolio in the position where it is most likely to achieve its intended principal objectives.

Figure 1.2 Purposes and content of portfolios¹



Obviously, a portfolio can be used to achieve more than one goal. When a portfolio is to serve a combination of goals, its position in the triangle will shift towards the centre as its strengths should be distributed more evenly over evidence, overviews and reflections. In practice, the majority of portfolios are not situated at one of the corners of the triangle.

A controversial issue in the literature on educational portfolios is whether it is acceptable to have one portfolio for both assessment and reflection.^{18,20,21} An argument against this dual function is that assessment may jeopardise the quality of reflection thereby detracting from the portfolio's effectiveness for mentoring purposes. Students may be reluctant to expose their less successful efforts at specific tasks and to reflect on strategies for addressing weaknesses if they believe they are at risk of having 'failures' turned against them in an assessment situation. Unassessed portfolios, on the other hand, do not 'reward' students for the time and energy they invest in them. As a result, students are likely to take the portfolio and any associated learning activities less seriously. We favour the middle ground and agree with the observation by Snyder and colleagues who wrote that: "The tension between assessment for support and assessment for high stakes decision making will never disappear. Still, that tension is constructively dealt with daily by teacher

educators throughout the nation”.²⁰ Striking the right balance between support and judgement is the challenge facing assessors/mentors with whom students talk about their portfolios.

Portfolios and learning from experience

Reflection

The central concept in constructivism and other modern views of education is that people learn by interpreting information, processes and experiences. These interpretations are the building blocks from which (a picture of) reality is constructed. Evidently, reality based on individual interpretations does not afford an objective view of reality but represents an individual, idiosyncratic view. It is this personal reality that guides a person’s actions. Reflection is a key concept in relation to the interpretation of reality. Despite considerable confusion about the precise definition of the term reflection,²²⁻²⁴ all authors writing about reflection share the view that human behaviour is guided by mental structures, structures that are not static but malleable and evolving and changing in response to experiences or confrontations with situations. Reflection can shape these mental structures.^{23,25} Based on this consensus view, Korthagen and colleagues defined reflection as the “mental process of trying to structure or restructure an experience, a problem, or existing knowledge or insights”.²³

Reflective learning

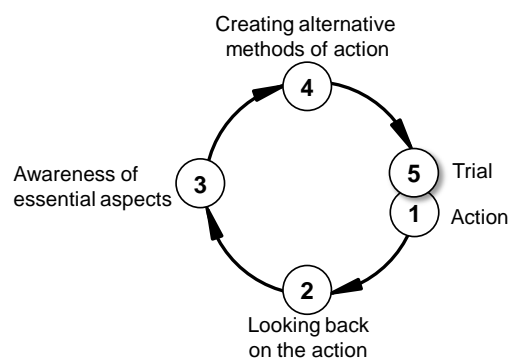
In that vein, reflective learning occurs when people “try to structure or restructure their internal representation of reality or the way they are using these representations for developing their plans or actions”.²³ Students learn from experiences in a cyclic process moving from acting to identification of strengths and weaknesses, of which the latter are then addressed through the pursuit of learning objectives formulated specially for that purpose, after which there is new action and the cycle repeats itself.

Based on their definition of reflection, Korthagen and colleagues developed an ideal model for cyclic professional development, based on reflection on experience. They refer to it as the ALACT model, named after the first letter of the five phases that are distinguished in the model (figure 1.3).²³

Starting with *Action* undertaken for a specific purpose, the ALACT cycle moves to the stage where a person looks back on a previous action, usually when it was not successful or something unexpected happened. This looking back is assumed to be accompanied by evaluation of whether the set goals

were reached. In many cases this may be regarded as self-assessment. Subsequently, awareness of crucial aspects revealed by the analysis of previous actions may trigger a search for alternative strategies for action, or, in the absence of alternative strategies, abandonment of the original goals. Finally, an alternative strategy is tested in a new action to see whether it is more effective in reaching the desired goal.

Figure 1.3. The ALACT model describes the phases of one cycle in the spiral of professional development²³



Learning to learn reflectively

Most students tend to find reflecting on their own learning uncongenial.¹⁴ For example, students arriving at university fresh from secondary education are not used to deliberately reflect on their learning. Outside the domain of medical education, there is substantial evidence that it takes considerable effort for students to learn to reflect on their actions or learning.²⁶

Korthagen et al. observed that some resistance to reflective learning was common among students on their first introduction to this approach.²³ Students' usually have had very different experiences at school, where transfer of knowledge was the focus of education. Thus students' images and expectations of education tend not to be in keeping with learning from reflection.²⁷ Korthagen et al. proposed a stepwise strategy to habituate students to learn from reflection.²³ Initially, the teacher sets clear assignments, offers clear choices and gives feedback and other support when students ask for help either tacitly or directly. As students become accustomed to the new approach and gain experience they are given more 'freedom'. At first, students are asked to reflect on limited, straightforward problems or experiences. As students become more comfortable with reflecting, they are asked to reflect on more complicated and critical experiences and problems. According to Korthagen et al., students can be

said to have become truly self-directed learners when they are able to move through the phases of the ALACT model on their own without help from a supervisor or mentor.²³

In the early stages of medical education, the emphasis is inevitably on the acquisition of knowledge and skills, although realistic and occasionally genuine cases are used as starting points for learning. Although the portfolio is especially suited to coaching and assessment of learning from experience at more senior levels of training,²⁴ it is considered appropriate to introduce a portfolio in the early years of the curriculum to help students develop reflective skills. For students of both the first and the final year of medical training, reflecting on their own learning appears to be not self-evident.¹⁴ Students arriving fresh from secondary education are not used to conscious reflection. Outside the domain of medical education, there is substantial evidence that students have to learn to reflect on their actions or learning.²⁶ Careful and extensive introduction of a portfolio in the early stages of the curriculum seems to be the way to provide for this training in reflection.²⁸⁻³⁰

As part of the assessment programme, the portfolio can be used to identify students who have difficulty in acquiring reflective skills. For these students extra mentoring may be an option.

The next paragraph gives the main research questions of the thesis.

Research questions

First analyses and design

The main argument for this thesis is to train self critical doctors by stimulating and assessing reflective skills with a portfolio in an early stage of the medical study. We tested a portfolio designed to develop and assess students' reflective ability. Central in the thesis is the portfolio as both a learning and an assessment method.

We adopted an approach resembling Korthagen et al.'s strategy of gradual and structured learning of reflective skills.²³ Although portfolios are particularly suitable for mentoring and assessment in workplace learning, our research was based on the assumption that a well-designed process centred round a portfolio in the first year of the undergraduate medical curriculum would be the best way of training students how to self reflect and of assessing students' accomplishments in. When we launched the project in 2001, many assumptions and beliefs about portfolios and reflection were circulating. However, a review of the literature quickly revealed a disappointing paucity of empirical data about portfolios aimed at stimulating and assessing medical students' reflective skills. Equally, we had no recourse to a portfolio model

from which we could derive the structure/organisation of the portfolio and guidelines for social interaction and assessment.³¹ This dearth of relevant practical and theoretical information concerning the type of portfolio and reflection we had in mind was reason for us to undertake the project reported in this thesis.

The project is set in year 1 of the undergraduate curriculum of Maastricht Medical School. The portfolio that is at the centre of this thesis was introduced as part of a curricular reform which went ahead in September 2001. As the new curriculum was still under development we had some freedom to design a portfolio based on insights from the literature. Nevertheless, the setting for which the portfolio was intended did somewhat constrain our liberty as designers. First of all, the portfolio had to be acceptable to students and teachers. Both groups were entirely unfamiliar with portfolios and we know only too well that unfamiliarity is bound to breed contempt.³⁰ Secondly, acceptability coupled with financial constraints forced us to create a portfolio that would make limited demands on teachers' time. Thirdly, year 1 of the new curriculum was largely devoted to problem-based, theoretical education in small groups, whereas the portfolios described in published research were mostly tailored to workplace learning or special projects.³¹ And last but not least, assessment of the portfolio was firmly scheduled for the end of year 1, because, for selection purposes, students were required to pass a summative end-of-year examination including assessment of all courses and components of year 1.

We sought to answer the questions:

1. *Does working with a portfolio support the development of reflective skills?*
2. *Can a portfolio be used to assess reflective skills?*

We did six studies to find answers to these questions. In each study different sub questions were addressed. At the end of this chapter we give an overview of the main and sub questions of this thesis

Chapter 2 describes a portfolio model to support the reflective teaching approach based on principles derived from the literature of medical education and other fields. The question addressed is:

What are the features of a portfolio that can support and assess the development of reflective skills?

It was a precondition that assessment of the portfolio would stimulate the development of students' reflective skills. This was our main point of departure in designing the assessment procedure. Earlier, we mentioned the ongoing debate in the literature about the feasibility and acceptability of combining acquisition and assessment of reflective skills in one portfolio. Opponents argue that students are not free to choose the content of their reflections when these are going to be assessed.

Chapter 3 shows how we managed to design an assessment procedure, using qualitative research strategies, which adhered to the principles of a reflective portfolio.

How can qualitative research strategies be used in a portfolio assessment procedure to ensure reliable and valid judgement?

This question refers to main question 2.

Evaluation

We used qualitative methods to explore which components of our portfolio were effective. We asked the mentors who supervised students' portfolios to share with us their views of the success of the portfolio enterprise and which factors, according to them, were decisive in this respect. Our decision to use teachers as a source of information was inspired by the literature where the role of the teacher is seen as crucial to the success or failure of portfolios.¹⁴

The central question of Chapter 4 is:

What are the conditions for successful reflective use of portfolios?

This question refers to main question 1.

We also used information from the mentor interviews to examine which criteria determine the quality of a portfolio. Based on the literature and data from the interviews we constructed an instrument to analyse the quality of a reflective portfolio. The instrument enabled us to examine two controversial key issues on which the literature did not yield any empirical data.

The first controversy is the validity of the assessment procedure. Undeniably, a holistic procedure holds promise as it allows for personal judgement to impact on ratings. Assessors have some freedom of judgement and have to make decisions about various aspects of portfolio quality.

A striking finding to emerge from our experiences with the Maastricht assessment procedure was the diversity in the way students presented their portfolios. There was tremendous variety in linguistic quality, in quality of lay-out and in the organisation and design of the portfolio. However, the assessment procedure was intended to appraise students' reflective skills only and not their skills in putting together a portfolio. This gave rise to concern

that a holistic assessment procedure might be subject to bias from presentation factors, such as writing style, design and lay-out (Chapter 5).

The questions:

Which criteria affect raters' judgements of students' reflective skills and Which criteria carry the most weight?

were addressed in a quantitative study in which we determined the quality of portfolios based on both form and content aspects and then correlated these outcomes with the actual assessments of the same portfolios which were done before we performed this study.

The second controversy, concerning electronic versus paper portfolios, is the subject of Chapter 6. Many authors assume that electronic portfolios offer considerable added value compared to paper ones. Among teachers and educational administrators and developers the prevailing preference is for electronic variants and there are even communities, personal web logs, conferences and lobbies solely devoted to promoting the electronic portfolio. However, there is a glaring contrast between the popularity of the electronic portfolio and the availability of empirical data on its presumed added value. The literature yields an abundance of assumptions and suggestions but is remarkably silent on empirical outcomes. The instrument we had constructed to analyse portfolio quality enabled us to conduct a quantitative study to answer the question:

What are the differential effects of a web-based portfolio versus a paper-based portfolio regarding portfolio quality, user friendliness, and student motivation?

Second analyses and design

In recent years, intriguing results have emerged from studies of portfolio usage in medical education. Study outcomes were richly varied or downright contradictory. For instance, studies in the late 1990s reported quite positive results for portfolios for GPs and GP registrars^{16,33}, but in the 21st century researchers came with mixed³⁴ or negative³⁵ results. We remain in the dark as to:

Why are portfolios effective or ineffective?

In order to tackle this question we performed a systematic review of the literature in Chapter 7.

The outcomes of the different studies underline the importance of the mentor for the effectiveness of working with a portfolio. Theories and teaching models from teacher education proved useful for the development of a

reflective teaching approach, which can be used by mentors in both the early study phase and in the clinical study phase. Chapter 8 introduces the question:

How can a mentor/teacher stimulate the development of reflective skills in medical education?

Chapter 9 synthesises the results of all the studies in this thesis, weighs up the strengths and weaknesses, and discusses the implications for portfolio design.

Overview of main and sub research questions

Main research question 1

- Does working with a portfolio support the development of reflective skills?

Sub questions

- What are the features of a portfolio that can support and assess the development of reflective skills?
- What are the conditions for successful reflective use of portfolios?
- What are the differential effects of a web-based portfolio versus a paper-based portfolio regarding portfolio quality, user friendliness, and student motivation?
- Why are portfolios effective or ineffective?
- How can a mentor/teacher stimulate the development of reflective skills in medical education?

Main research question 2

- Can a portfolio be used to assess reflective skills?

Sub questions

- What are the features of a portfolio that can support and assess the development of reflective skills?
- How can qualitative research strategies be used in a portfolio assessment procedure to ensure reliable and valid judgement?
- Which criteria affect raters' judgements of students' reflective skills and which criteria carry the most weight?
- Why are portfolios effective or ineffective?

The research project that is the topic of this thesis was conceived, designed, conducted and described by a team. That is why throughout the thesis we write from a plural perspective. The research team consists of the following members: Erik Driessen, Jan van Tartwijk, Jan Vermunt and Cees van der Vleuten, with (linguistic) support from Mereke Gorsira. For different studies the team was reinforced by: Alice Drenthe, Arno Muijtjens, Elise van der Zijp, Ester Coolen, Floris van Blankestijn, Jonne van der Zwet, Karlijn Overeem, Lambert Schuwirth, Ron Hogenboom, Tim Dornan, Val Wass and Victor Burger. Obviously, the sole responsibility for the project rests with the PhD candidate, Erik Driessen.

References

- 1 Bleker O, Hoorntje J, Schelfhout V. Beter en leuker: CCMS ontvouwt plannen voor de vervolgopleiding van medisch specialisten. *Medisch Contact* 2004;59 (43):1692-5.
- 2 Smith R. All changed, changed utterly. *BMJ* 1998;316:1917-8
- 3 The Shipman Inquiry www.the-shipman-inquiry.org.uk
- 4 General Medical Council. *Revalidating Doctors: Ensuring Standards, Securing the Future*. London: GMC 2000
- 5 Royal College of General Practitioners. *Portfolio-based learning in general practice: Report of a working group on higher professional education* 1993. Occasional paper 63.
- 6 Zeiger RF. Toward continuous medical education. *J Gen Int Med* 2005;20(1):91-4.
- 7 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65(9 Suppl), S63-7.
- 8 van Tartwijk J, Driessen EW, Hoeberigs B, Kösters J, Ritzen M, Stokking K. *Using an electronic portfolio (in Dutch: Werken met een elektronisch portfolio)*. Groningen: Wolters Noordhoff 2003.
- 9 Shulman LS. Teacher portfolios: A theoretical activity. In Lyons N, eds. *With Portfolio in Hand: Validating the new teacher professionalism*. New York: Teachers College Press 1998.
- 10 Lyons N. Introduction. In Lyons N, eds. *With Portfolio in Hand: Validating the New Teacher Professionalism*. New York: Teacher College Press 1998.
- 11 Bird T. The schoolteacher's portfolio: an essay on possibilities. In Millman J, Darling-Hammond I, eds. *The new handbook of teacher*

- evaluation: Assessing elementary and secondary school teachers.* Newbury Park, CA: Corwin Press, inc 1990:241-256.
- 12 Järvinen A, Kohonen V. Promoting professional development in higher education through portfolio assessment. *Assessment and Evaluation in Higher Education* 1995;20(1):25-36.
 - 13 Snadden D, Challis M, Thomas M. *Portfolio-based Learning and Assessment: The use of portfolio-based learning in medical education* AMEE guide No. 11. Dundee: University of Dundee 1999.
 - 14 Wade RC, Yarbrough DB. Portfolios: a tool for reflective thinking in teacher education. *Teaching and Teacher Education* 1996;12(1): 63-79.
 - 15 Seldin P. *The Teaching Portfolio: A practical guide to improved performance and tenure/promotion decisions (2nd ed.)*. Bolton MA: Anker 1997.
 - 16 Mathers NJ, Challis MC, Howe AC, Field NJ. Portfolios in continuing medical education - effective and efficient? *Med Educ* 1999;33(7):521-30.
 - 17 Oermann MH. Developing a professional portfolio in nursing. *Orthopaedic Nursing* 2002;21(2):73-78.
 - 18 Wolf K, Dietz M. Teaching portfolios: Purposes and possibilities. *Teacher Education Quarterly* 1998;25(1):9-22.
 - 19 Spandel V. Reflections on portfolios. In Phye GD, eds. *Handbook of Academic Learning: Construction of knowledge*. San Diego: Academic Press 1997;573-91.
 - 20 Snyder J, Lippincott A, Bower D. The inherent tensions in the multiple uses of portfolios in teacher education. *Teacher Education Quarterly* 1998;25(1):45-60.
 - 21 Tigelaar DEH, Dolmans DHJM, Wolfhagen HAP, van der Vleuten CPM. Using a conceptual framework and the opinion of portfolio experts to develop a teaching portfolio prototype. *Studies in Educational Evaluation* 2004;30(4):305-21.
 - 22 Hatton N, Smith D. Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education* 1995;11(1):33-49.
 - 23 Korthagen FAJ, Kessels J, Koster B, Lagerwerf B, Wubbels T. *Linking theory and practice: The pedagogy of realistic teacher education*. Mahwah, NY: Lawrence Erlbaum Associates 2001.
 - 24 Mansvelder-Longayroux DD. *The learning portfolio as a tool for stimulating reflection by student teachers*. Leiden: Leiden University 2006.

- 25 Korthagen, F. A. J., & Wubbels, T. (1995). Characteristics of reflective practitioners: Towards and operationalization of the concept of reflection. *Teachers and Teaching: Theory and Practice*, 1(1), 51-72
- 26 Ertmer, PA Newby TJ. The expert learner: strategic, self-regulated, and reflective. *Instructional Science* 1996;24:1-24.
- 27 Weber S, Mitchell C. *That's funny, you don't look like a teacher: Interrogating images and identity in popular culture*. London:The Falmer Press 1995.
- 28 Loughran J, Corrigan D. Teaching portfolios: a strategy for developing learning and teaching in preservice education. *Teaching & Teacher Education* 1995;11(6):565-577.
- 29 Krause S. Portfolios in Teacher Education: Effects of Instruction on Preservice Teachers' Early Comprehension of the Portfolio Process. *Journal of Teacher Education* 1996;47(2):130-38.
- 30 Tartwijk J, Rijswijk P, Tuithof H, Driessen E. Using an analogy in the introduction and Construction of a portfolio. *Teaching and Teacher Education* 2008;24(4): 927-38.
- 31 Zeichner K, Wray S. The teaching portfolio in U.S. Teacher education programs: What we know and what we need to know. *Teaching and Teacher Education* 2001;17(5):613-21.
- 32 Finlay IG, Maughan TS, Webster DJ. A randomized controlled study of portfolio learning in undergraduate cancer education. *Med Educ* 1998;32(2):172-176.
- 33 Snadden D, Thomas ML. Portfolio learning in general practice vocational training - does it work? *Med Educ* 1998;32(4):401-6.
- 34 Dagley V, Berrington B. Learning from an evaluation of an electronic portfolio to support general practitioners' personal development planning, appraisal and revalidation. *Education for Primary Care* 2005;16:567-74.
- 35 Pearson DJ, Heywood P. Portfolio use in general practice vocational training: A survey of GP registrars. *Med Educ* 2004;38(1):87-95.

2 Use of portfolios in early undergraduate medical training

**Erik W Driessen,
Jan van Tartwijk,
Jan D Vermunt
Cees PM van der Vleuten**

Published in: Medical Teacher 2003; 25: 14-9

Summary

The ability to reflect on one's own action is seen as an important skill for a doctor. A thorough introduction of the portfolio planned in the early stages of their studies seems to be the way to train medical students in reflection. This article describes the use of portfolios in early undergraduate medical training. The literature on portfolios suggests 3 aspects that are crucial for the effectiveness of portfolios: structure, coaching and assessment. The portfolio system was designed by transposing the experience with portfolio systems outside and inside medical training to a situation of Year 1 medical students. During academic year 2001-02 242 Year 1 medical students compiled a portfolio. Student experience was collected by semi-structured interviews. The majority of students were of the opinion that analysing one's competences in a portfolio was instructive and meaningful. With regard to learning how to reflect and recognize learning needs, however, mentor coaching proved to be necessary. The results thus far show that the portfolio is a worthwhile addition to existing assessment and learning tools.

Introduction

Traditionally, the study of medicine comprises several years of theoretical learning followed by 1 or 2 years of practical training during clerkships. A first problem with this separation between studying theory and learning in practice is that, when students are deemed to perform the task of an independent doctor, they prove to be overwhelmed by the complexity of 'real-life' situations. Their first priority will be to cope with these situations. They will accept strategies that will help them to survive, even if they are incompatible with their insights from the first part of the curriculum. Moreover, particularly during the practical period, it often becomes evident that the students' theoretical knowledge is structured the wrong way: a patient presents him/herself with complaints and symptoms, whereas the students' knowledge is structured around diagnoses.¹ Clerks are often almost literally told: "Forget theory: in practice we do things this way." Thus, they develop practical knowledge and strategies that will guide their performance as doctors, suppressing the theoretical knowledge acquired in the earlier stages of the study.

The gap between theory and practice causes an increasing number of medical schools to revise their curricula in order to teach their students to function in authentic professional situations at an earlier stage of their studies. The combination of authentic learning and theory instruction should better enable students to relate theory to practice.

An important tool for coaching and assessing authentic learning is the portfolio. In a portfolio students look back on their actions systematically, analyse them with the aid of, among other things, theoretical knowledge and designate alternative methods of action. We will refer to such a form of systematic self-regulation as "reflection".² In their portfolios students underpin their reflections with illustrative materials.

The literature reports the use of portfolios in several educational contexts, including clerkships in undergraduate medical training, postgraduate training and continuous professional development. Snadden and Thomas describe how in general practice vocational training the portfolio serves as a tool for reminding, planning, tracking and encouraging reflection.³ Pitts *et al.* concluded that portfolios are a valuable source for learning in general practice but their use as an assessment tool may be more problematic.⁴ Challis reviewed the role of portfolios as a tool for assessment.⁵ Another example of the use of portfolios is the undergraduate cancer projects at the University of Wales College of Medicine.⁶ Compiling portfolios enables the students to reflect on a number of emotional and communicative aspects of

working with cancer patients throughout the project. At the Medical School of Dundee, portfolios are used for summative assessment of final-year students.⁷ These and other examples are aimed at those phases of medical training that are characterized by a large degree of authentic learning in the workplace.

In the new Maastricht curriculum, authentic learning is already introduced in the first year (<http://www.fdg.unimaas.nl/bib/curriculum2001>). Naturally, the emphasis is on acquisition of knowledge and skills, but authentic and sometimes genuine cases are used as a starting point. Although the portfolio is especially suited for coaching and assessing authentic learning at more senior levels of training, it was felt appropriate to introduce the portfolio in the early years of the curriculum. The express purpose was to develop the students' reflective ability. For students of both the first and the final year of medical training, reflecting on their own learning is not self-evident.⁸ Students arriving fresh from secondary education are not used to conscious reflection. Outside the domain of medical education, there is substantial evidence showing that students have to learn to reflect on their actions or learning.⁹ A thorough introduction of the portfolio, planned in the early stages of their studies, seems to be the way to provide for this training in reflection.^{10,11}

In addition to stimulating reflection, we had 3 other reasons for an early introduction of portfolio in our curriculum. The second reason was to create a mentoring system in which student could receive pastoral care during their studies. A third was to give students more responsibility for their learning and assessment. The last goal was to make the portfolio part of the assessment system, forcing students to review and integrate all assessment feedback.

In this article, we will describe the Maastricht portfolio as an illustration of an early undergraduate portfolio. We will discuss the reasons for the choices we made in designing the portfolio and report on a first evaluation, focusing on the effects on students' reflective ability.

The Maastricht Portfolio in the First Year of Medical Studies

In the period (March 2000 until August 2001) in which we designed the portfolio system, it was not possible to draw on the experience from other programmes with similar portfolio systems. Both Medline and Eric were searched for empirical studies of portfolios. Medline produced studies in the senior years of undergraduate and in postgraduate training and professional nurse training. There are also practical manuals available for the application

of the portfolio learning and assessment tool in the final phase of medical studies and for adult education.^{12,13}

Not 1 study found related to Year 1 medical students. The search in Eric did produce empirical studies of portfolios in the early years of instruction: in most cases they were deployed in teacher training and pre-university education. In addition, Eric yielded studies relating to other phases of non-medical education, of course. In designing our portfolio system, we had therefore to transpose the experience with portfolio systems outside and within medical training to a situation of Year 1 medical students. Below, we will review a number of essential aspects of the portfolio and discuss the reasons for our choices on the basis of evidence found elsewhere.

Structuring

Many portfolio systems prove to have too much of a prescribed character: the exact content, the volume and the structure of the portfolio have been pre-set in detail.¹⁴ In compiling such a portfolio, students have little or no possibility to bring out their personal qualities. An overly structured portfolio will readily deteriorate into a tick-off list through which students can demonstrate that they have met the expectations. This results in students showing forced study behaviour and being mainly occupied with getting the learning results needed ticked off.^{15,16} In addition, it has become evident that, if students are afforded little freedom compiling the portfolio, they will not readily include their observations.¹⁷ Student ownership is an important contributor to successful stimulation of reflection.

If, however, little or no structure is provided for the portfolio, students often have no idea how to go about it.^{8,16} Portfolio systems that provide too little structure, often lead to frustrating experiences, especially for students who are compiling a portfolio for the first time.

For the Maastricht Year 1 portfolio a compromise was sought. Since the aim of the portfolio is to encourage reflection, it is advisable to organize the portfolio around student self-reflection.¹⁶ In the portfolio reflection is structured around 4 professional roles of the doctor, based on a Dutch national competence profile for doctors: a) one's role as medical expert; b) researcher, c) healthcare worker; and d) person.¹⁸ Professional roles were chosen to express competence, because these proved to be easier to comprehend than abstractly formulated competences.¹⁹ The portfolio has been given structure by including sections, 1 for each professional role. Self-assessment and analysis of the student's own performance is thus organised on a role-by-role basis, underpinned by systematic reference to illustrative materials.

The structure of the portfolio is as follows:

- curriculum vitae
- role as medical expert;
- role as researcher;
- role as health care worker;
- role as person;
- general;
- summary of strengths and weaknesses analysis;
- report of progress interview and/or of exit interview/ recommendation, and
- Annexes.

Global requirements are set for each professional role. The students themselves determine the content and form of their portfolio. Their self-reflection serves both them when compiling the portfolio and the persons reading it as a guideline. The students are free in their choice of materials. For each professional role, a number of probing questions has been formulated to guide the student in writing the introspective part. Posing probing questions is a way to encourage reflection on learning.²⁰ The questions structure the students' thought processes.

For each professional role, the students analyse their professional development thus far, draw conclusions from the analysis and set learning objectives for the coming period. They therefore show per professional role what and how they have learned and how they wish to improve.

To aid the students in this, a manual is provided in which each role is explained. The manual also contains the global objectives of the year, possible questions the students may ask themselves and potential sources of information that may be used to document the advancement in the areas of competence in question, illustrated by examples.

To illustrate this Box 1 shows the requirements, probing questions and sources of information for the role as health care worker (Appendix 1)

Box 2 contains a sample fragment from a portfolio from an individual student (Appendix 2).

Mentoring

A crucial factor for the effective use of portfolios aimed at stimulating reflection, is regular discussion of the portfolio with others.^{21,22} Both in one-on-one contacts between the student and the mentor and by students reading and providing comment on each other's portfolios.^{3,23} The Royal College of General Practitioners advocates cooperation between the learner and the

mentor in working with portfolios for three reasons.²⁴ In the first place, the mentor assists the students in recognizing their learning needs and setting up a learning schedule. Learners are often inclined to formulate learning schedules for aspects they have already mastered and ignore areas in which there remains a lot to be learned. In the second place, the mentor ensures that not only the practical but also the emotional aspects of working with patients become part of the learning process. The emotional part of working with patients is often neglected. A third reason for working with mentors is that the mentor is able to validate the portfolio materials. As a result of the personal relation between mentor and students, the mentor is able to appreciate the value of the evidence.

We opted for a mentoring system. A single mentor provides support for twenty students in developing their portfolios and provides general pastoral guidance. At least twice a year, student and mentor conduct an individual interview; a progress interview and an end-of-year interview. During these interviews, the student's performance in each area of competence is evaluated, strong and weak points are identified and new learning objectives formulated for the coming period. In so doing, the students can better identify the focus of their studies and are able to actively steer their learning. Thus, the students develop skills that are relevant to their subsequent studies and future profession, such as self-reflection and a critical and focused learning attitude.

For the end-of-year assessment, the mentor does not assess his/her own students' portfolios. He/she only serves as coach and advisor.

The mentor's responsibility goes beyond supervising the portfolio: he/she also has a general pastoral function. If a student has a problem or a question that has consequences for his/her studies, he/she may direct himself to his mentor. If he/she has problems of a structural nature, the mentor can refer the student to a student counsellor or other, specialized professionals.

Assessment

The assessment literature has shown that assessment drives learning and that congruence between training and assessment is imperative.²⁵ Assessment can also be used to steer student learning in a desirable direction. Having students complete portfolios without some form of reward (credits) was considered to be ineffective in the long run. On the other hand, portfolio completion was not to become a "test-achievement" ritual for the purpose of passing. For this reason, we opted for the portfolio as a modest part of the

entire assessment program, which needed to be carried out properly, however, in order for the student to be promoted to the next year.

Can 2 different goals -assessment and reflection – be combined in a single portfolio? There is a possibility that students choose not to show their weaknesses if the portfolio is also assessed. In the literature, this is sometimes referred to as “the corruption of portfolios for testing purposes.”²⁶ Such corruption is often the result of strategies employed to improve the reliability of portfolio assessment. Research into the quality of portfolio assessment shows that the reliability of the assessment is one of the weaker aspects of portfolio use in education.²⁷

Reliability, perceived of as inter-rater reliability, may be improved by three strategies: portfolio standardization; assessment objectification by using analytical criteria; and increasing the number of assessors. The first 2 strategies are at right angles with the philosophy underlying the portfolio. Originally, portfolios were introduced to make more qualitative decisions on the basis of authentic materials coupled to unique personal experience.²⁸ Portfolio standardization and the use of analytical criteria with the aim of improving reliability will threaten validity, because it limits the room for describing students’ personal learning experiences in different authentic situations. The third strategy for enhancing inter-rater reliability, i.e. increasing the number of assessors, is an effective strategy in theory. Even low inter-rater reliability may be compensated by involving more assessors. However, for most educational institutions this strategy is too expensive, especially if large numbers of students need to be assessed.

In assessing portfolios, we cannot avoid a more holistic and qualitative approach requiring a more “professional judgement” on the part of the assessor.²⁹ However, some degree of standardisation may be introduced. Of relevance here is the credibility or justifiability of the assessment procedure. It must be organized in such a way that assessment at the end of the procedure does not present surprises for the student. A system of checks and balances must be in place enabling the student to work towards an anticipated end result.

To attain this, we introduced multiple portfolio discussion moments in the assessment procedure: progress interviews and exit interviews between mentor and student, followed by a separate assessment of the portfolio by a committee. If a mentor is not satisfied with the quality of the portfolio, he/she brings this up during the interviews. Those aspects that did not work out too well, may be improved before the final assessment. By so doing, the student receives intermediate feedback on the quality of his/her portfolio. By

including this type of feedback loops into the portfolio procedure, in effect, a combination of information is achieved by building in several moments of measurement. If this procedure is followed correctly, assessment by the assessment committee at the end of the year will not lead to unexpected outcomes. The committee has access to the documented intermediate evaluations and can take these into account in its assessment. A drawback is that, as a result of this, the committee's assessment is not fully independent. The assessment procedure is organized sequentially: if it is difficult to determine the quality of the portfolio, more than 1 assessor from the assessment committee is involved in the procedure. If mentor, student and/or assessor disagree on the assessment of the performance shown in the portfolio, an additional assessor will be involved. Subsequently, the complete assessment committee discusses the portfolio. If at the start of the procedure mentor, student and assessor agree on the quality of the performance shown in the portfolio, only 1 member of the assessment committee reads and grades the portfolio: "insufficient", "sufficient" or "with merit".

A first evaluation

Two hundred and forty-seven Year 1 students compiled a portfolio during the academic year 2001-02. Semi-structured interviews were held with a select group of students to explore the effect of the portfolio on reflection. Students from 4 random mentor groups, 39 in total, were approached and asked if they would be interviewed on the portfolio. With the exception of 1 student, all students were prepared to take part in the interview. The interview schedule focused on the process of compiling a portfolio and its possible effect on reflective ability. In addition to this, the students were given the opportunity to offer suggestions for improvement.

All students reported that they were capable of compiling a portfolio in the initial stage of their studies. A typical remark was:

"I think it is okay that we start with it (portfolio) so early, because then you can get used to it" (student 5)

The students had a positive opinion of the portfolio acting as an incentive to reflection on their progress. Indeed 95% of the students interviewed indicated that the portfolio was an important stimulus to carry out strengths and weaknesses analyses. Typical remarks were:

"Because you have to write down your strong and weak points you are forced to think about yourself" (student 1)

"Because there are different professional roles (in the portfolio), you approach yourself from different perspectives" (student 5)

"Normally I put a test on which I scored badly aside...but if you put all your illustrative materials next to each other, I can exactly see how I work" (student 6)

In all, 80% percent of the students regarded the strength and weaknesses analyses as helpful.

"It is difficult to think about your own functioning, but it is surely useful."
(Student 1)

"... the self-analyses made me think about myself." (Student 38)

The portfolio therefore encouraged them to work systematically on readjusting their actions: 87% observed that the portfolio had helped them to formulate learning objectives. Typical remarks were:

"... I found out what I did wrong, and by formulating learning objectives, I started to study in a different way to achieve what I want to achieve"
(student 3)

"... because you have to verbalize it and read it over, you can better find connections and better draw conclusions and thus formulate learning goals" (student 14)

The structure of the portfolio proved to meet expectations: on the one hand, it gave the students a clear picture of what was expected of them, whereas it also gave them freedom to personalize their portfolio.

"You have much freedom in what you write down (...). If you just say to somebody: "make a portfolio", then that person has no structure and does not know how to start. This division into 4 roles provided something to hold on to." (student 9)

One of the professional roles – that of healthcare worker- proved difficult to perform in practice, because the students had little relevant experience in this area. Some students had resolved this problem by including pre-university experience and experience as patients or as relatives of patients in this section of the portfolio.

Conclusion

The main reason for the introduction of portfolios in early undergraduate medical training was to develop the reflective ability of early undergraduate

students. Reflective ability is seen as an important skill to learn from practice and for life-long learning. The ability to reflect can protect students from the shock of practice when entering their clerkships. The literature suggests that a thoroughly planned introduction in the early stages of study is an effective way for students to learn how to reflect. The first experiences with the portfolio bear out the view that an early introduction of the portfolio in medical training is an effective tool to stimulate students' reflective ability. The portfolio seems to act as an incentive to reflection on behaviour. It encourages students to work systematically on readjusting their actions.

We used experience with portfolio systems outside and within medical training to develop a portfolio for Year 1 medical students. Literature on portfolios suggests that the factors contributing to portfolio effectiveness include a supportive mentor system, clear portfolio structure, an appropriate assessment procedure and early and unambiguous portfolio introduction. The first experiences with our portfolio suggest that these factors are indeed crucial. The majority of students were of the opinion that analysing one's competences in a portfolio was both instructive and meaningful. With regard to learning how to reflect, however, the mentors felt that coaching by a teacher proved to be necessary. The mentoring system may be a major explanation for the fact that students see the portfolio as effective in stimulating reflective ability. Other studies of the effect of portfolios on reflective ability often show disappointing results.^{8,27} In many cases, the purpose of portfolios seems unclear to students. As a result, portfolios are often superficial. In other words, students have to learn how to reflect and compile a portfolio. Teacher support is essential to this learning process.

In conclusion, we feel that a portfolio constructed according to the considerations described in this paper is a worthwhile addition to existing assessment and learning tools. Our preliminary findings show that the introduction of such a portfolio in the early stages of medical training seems to be an effective way for students to learn how to reflect on their learning and behaviour. Reflective skills are essential to both learning from practice and lifelong learning. It is our hope, therefore, that early introduction of the portfolio will help develop the appropriate skills for more in-depth portfolio learning and assessment in later stages.

References

- 1 Prince KJAH, Wiel MWJ, Scherpbier AJJA, Vleuten CPM, Boshuizen HPA. A Qualitative Analysis of the Transition from Theory to Practice in Undergraduate Training in a PBL-Medical School. *Adv Health Sci Educ* 2000;5:105-116.
- 2 Korthagen F. *Linking practice and theory. The pedagogy of realistic teacher education*. Mahwah (NJ): Lawrence Erlbaum Associates; 2001.
- 3 Snadden D, Thomas ML. Portfolio learning in general practice vocational training - does it work? *MedEduc* 1998;32:401-406.
- 4 Pitts J, Coles C, Thomas P. Enhancing reliability in portfolio assessment: 'shaping' the portfolio. *Med Teacher* 2001;23(4):351-355.
- 5 Challis M. Portfolios and assessment: meeting the challenge. *Med Teacher* 2001;23,(5):437-440.
- 6 Finlay IG, Maughan TS, Webster JT. A randomized controlled study of portfolio learning in undergraduate cancer education. *Med Educ* 1998;32(2):172-176.
- 7 Davis MH, Friedman BDM, Harden RM, Howie P, Ker J, McGhee C, Pippard MJ, Snadden D. Portfolio assessment in medical students' final examinations. *Med Teacher* 2001;23:357-366.
- 8 Wade RC, Yarbrough DB. Portfolios: a tool for reflective thinking in teacher education? *Teaching & Teacher Education* 1996;12(1):63-79.
- 9 Ertmer PA, Newby TJ. The expert learner: strategic, self-regulated, and reflective. *Instructional Science* 1996;24:1-24.
- 10 Loughran J, Corrigan D. Teaching portfolios: a strategy for developing learning and teaching in preservice education. *Teaching & Teacher Education* 1995;11(6):565-577.
- 11 Krause S. Portfolios in Teacher Education: Effects of Instruction on Preservice Teachers' Early Comprehension of the Portfolio Process. *Journal of Teacher Education* 1996;47(2):130-138.
- 12 Snadden D, Thomas ML, Challis M. *AMEE Medical Education Guide No. 11 (revised): The use of portfolio-based learning in medical education*. Dundee: Association for Medical Education in Europe 1999.
- 13 Friedman BDM, Davis MH, Harden RM, Howie PW, Ker J, Pippard MJ. AMEE Medical Education Guide No. 24: Portfolios as a method of student assessment. *Medical Teacher* 2001;23(6):535-551.
- 14 Simmons J. Control the purpose, not the contents: coaching the creation of teaching portfolios. *Action in teacher education* 1996;18(1):71-81.

- 15 Carrol JA, Potthoff D, Huber T. Learning From Three Years of Portfolio Use in Teacher Education. *Journal of Teacher Education* 1996;47(4):253-262.
- 16 Gearhart MG, Wolf SA. Issues in Portfolio Assessment: Assessing Writing Processes from their Products. *Educational Assessment* 1997;4(4):265-296.
- 17 Murphy S, Bergamini J, Rooney P. The Impact of Large-Scale Portfolio Assessment Programs on Classroom Practice: Case Studies of the New Standards Field-Trial Portfolio. *Educational Assessment* 1997;4(4):297-233.
- 18 Metz JCM, Verbeek-Weel AMM, Huisjes HJ. *Raamplan 2001 artsopleiding (Blueprint 1994: training of doctors in the Netherlands)*. Nijmegen: Mediagroep Nijmegen 2001.
- 19 Tartwijk J, Lockhorst D, Tuithof H. Universiteit Utrecht: Portfolio's en de opleiding van docenten (Utrecht University: Portfolios and teacher education) In Driessen EW, Beijaard D, van Tartwijk J, van der Vleuten CPM. eds *Portfolios*. Groningen:Wolters-Noordhoff 2002:75-88.
- 20 Spandel V. Reflections on portfolios. In Phye GD, eds. *Handbook of Academic Learning: Construction of knowledge*. San Diego: Academic Press 1997;573-91.
- 21 Golberg GL, Roswell BS, Michaels H. Can Assessment Mirror Instruction? A look at Peer Response and Revision in a Large-Scale Writing Test. *Educational Assessment* 1996;3(4):287-314.
- 22 Wolf K, Whinery B, Hagerty P. Teaching portfolios and portfolio conversations for teacher educators and teachers. *Action in teacher education*. 1995;17(1):30-39.
- 23 Freidus H. Mentoring portfolio development. In Lyons N. eds *With portfolio in hand: validating the new teacher professionalism*. New York:Teacher College Press 1998:51-68.
- 24 Royal College of General Practitioners. *Portfolio-based learning in general practice: Report of a working group on higher professional education* 1993. Occasional paper 63.
- 25 Driessen EW, Vleuten CPM. Matching student assessment to problem-based learning: lessons from experience in a law faculty. *Studies in Continuing Education* 2000;22(2):235-248.
- 26 Huot B. Beyond the classroom: using portfolios to assess writing. In Black L, Daiker DA, Sommers O, Stygall G. eds *New directions in portfolio assessment: Reflection, practice, critical theory and large-scale scoring*. Portsmouth:Cook Publishers 1994:325-333.

- 27 Beijaard D, Driessen EW, Tartwijk J, Vleuten CPM Portfolio's verder uitgewerkt (Further development of portfolios), In Driessen EW, Beijaard D, van Tartwijk J, van der Vleuten CPM. eds *Portfolios*. Groningen: Wolters-Noordhoff 2002:119-136.
- 28 Shulman LS. Teacher portfolios: A theoretical activity. In Lyons N, eds. *With Portfolio in Hand: Validating the new teacher professionalism*. New York: Teachers College Press 1998.
- 29 Norman GR, Vleuten CPM, Graaff ED. Pitfalls in the pursuit of objectivity: issues of validity, efficiency and acceptability. *Med Educ* 1991;25:119-126.

Appendix 1

The Doctor as Health Care Worker

First-Year Portfolio Terms:

Becoming acquainted with health care practice; Being able to determine and express your position when discussing ethical problems in medical practice.

Probing questions for strengths and weaknesses analysis:

(You need not answer all questions literally in your strengths/weaknesses analysis.)

What is your experience as a health care worker?; In what way is your experience different from your expectations?; Has your idea of working in health care changed?

What was your experience in working together with other health care professionals?; How do they deal with one another?

What was your experience in discussing ethical conduct in medical practice with fellow students?; Was it difficult to get your position across?; Did your position on a particular ethical issue change as a result of the discussion?

Did you notice differences in the care systems within the health care sector?

Possible sources of information:

Practice reports and practice experience; professional behaviour assessment in practice and instruction; feedback from Medical Practice Supervisor (MPO); possible experience during on-the-side job as health care worker; instruction taken at other universities, and so on.

Appendix 2

Sample Fragment:

Strengths and weaknesses analysis fragment

“I had known for a long time that I wanted to study medicine. This has been the main reason why I took an on-the-side job in the care sector (see resume). This way, you acquaint yourself already with health care and you are able to gain a better idea of your future profession. Contributing to the wellbeing of patients appeals to me, but at times you experience less pleasant things. Each week, you provide care to people and as a result you establish some kind of relationship. It is therefore not so easy when someone passes away.

My idea of health care has not changed much since I took up medicine. What did strike me was that as a medical student during a traineeship, you are suddenly assigned a different role than when you were an assistant carer. You do not belong, in fact, not with the nurses and also not with the doctors. I really had to get used to that. When discussing ethical questions with other students, I can get really worked up. Some students are quite opinionated when discussing euthanasia or abortion. It should all be possible. I feel that when you take a good look at the health care sector, you will get a more balanced view. Because I get irritated about other students' remarks, I find it harder to defend my own position (see feedback MPO-supervisor, annex no. 6). I find it much easier to put my position in writing (see feedback report, annex no. 7).”

Learning objective: - To be able to properly express my views on ethical questions in discussions with my fellow students

3 The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study

**Erik W Driessen
Cees PM van der Vleuten
Lambert Schuwirth
Jan van Tartwijk
Jan D Vermunt**

Published in: Medical Education 2005; 39: 214-220

Summary

Aim

Because it deals with qualitative information, portfolio assessment inevitably involves some degree of subjectivity. The use of stricter assessment criteria or more structured and prescribed content would improve inter-rater reliability, but would obliterate the essence of portfolio assessment in terms of flexibility, personal orientation and authenticity. We resolved this dilemma by using qualitative research criteria as opposed to reliability in the evaluation of portfolio assessment.

Methodology/research design

Five qualitative research strategies were used to achieve credibility and dependability of assessment: triangulation, prolonged engagement, member checking, audit trail and dependability audit. Mentors read portfolios at least twice during the year, providing feedback and guidance (prolonged engagement). Their recommendation for the end-of-year grade was discussed with the student (member checking) and submitted to a member of the portfolio committee. Information from different sources was combined (triangulation). Portfolios causing persistent disagreement were submitted to the full portfolio assessment committee. Quality assurance procedures with external auditors were used (dependability audit) and the assessment process was thoroughly documented (audit trail).

Results

A total of 233 portfolios were assessed. Students and mentor disagreed on 7 (3%) portfolios and 9 portfolios were submitted to the full committee. The final decision on 29 (12%) portfolios differed from the mentor's recommendation.

Conclusion

We think we have devised an assessment procedure that safeguards the characteristics of portfolio assessment with credibility and dependability of assessment built into the judgement procedure. Further support for credibility and dependability might be sought by means of a study involving different assessment committees.

Introduction

The use of portfolios as an assessment method has gained rapid popularity. As has happened with many assessment instruments, the term “portfolio” has become a container concept covering a diversity of methods.^{1,2} At the heart of every portfolio is information collected in evidence of the owner's learning process and/or competence levels. The evidence is often organised by competencies and may be supplemented with reflections on educational achievement and personal and professional development.³ Portfolios were primarily introduced to assess performance in authentic contexts and encourage learners to reflect on their performance.⁴ When portfolios are used for summative rather than formative assessment, the psychometric qualities must meet stringent requirements, particularly in terms of reliability. What evidence the scant studies on the reliability of portfolios have revealed is cause for concern, as most studies reported moderate to low (inter-rater) reliability.^{3,5,6} The inevitable conclusion is that extreme caution is warranted when portfolios are used for summative purposes.⁷ A case in point is the study carried out in the state of Vermont on the reliability of their large-scale portfolio assessment programme in primary education.⁸ In response to the low reliability scores (inter-rater (Spearman) correlations between 0.45 and 0.65), the Vermont Department of Education restricted the reporting of portfolio scores. The Vermont case attracted substantial attention and led to increased standardisation of portfolio assessment.

There are several strategies for improving inter-rater reliability in portfolio assessment:

1. standardisation, for example by structuring content and restricting the number of admissible sources of evidence;
2. rater training and the structuring of judgement through checklists with strict performance criteria, and
3. using large numbers of raters to average out any rater effects.

The disadvantage of the first 2 strategies is that they jeopardise validity. Portfolios are valuable largely because of the richness of the information they supply. They enable students to present documentation of their personal, authentic, educational experiences and experiences in real practice. Standardising those experiences would inevitably detract from the portfolio's educational value. Training of raters and shared rater experiences would

improve inter-rater reliability, but studies on the OSCE have warned against unrealistically high expectations of rater training, even with well-defined instruments.⁹ The lesson that detailed checklists can easily trivialise assessment has been learned in other assessment domains as well.¹⁰ Increasing the number of raters would be an effective strategy, were it not for practical constraints, such as the time-consuming nature of portfolio judgement. In summary, portfolio assessment appears to be caught between the 2 classic evils of poor reliability and poor validity. This begs the question of how to achieve sufficient reliability of qualitative and subjective judgements for summative purposes without falling into the trap of “corruption of portfolios for testing purposes”.¹¹

Common misconceptions pervading this discussion are that subjectivity equals unreliability and that objectivity equals reliability. This is not universally true: objective examinations may be unreliable (cf. a single-item, multiple-choice examination) and - more importantly - subjective judgements can be reliable provided an adequate number of different judgements are collected and collated.¹² In any formal assessment procedure a fair decision must optimally reflect the demonstrated competence. This implies that assessments must be comparable across candidates, with minimisation of bias and error. Psychometrically, this means large sample sizes and structured (i.e. objective) assessments. In this paper we will present a qualitative approach to portfolio assessment that can enhance reliability without taking recourse to large samples and rigid structuring. Before looking at reliability from the perspective of qualitative research criteria, we will address some analogies between concepts. Some of the concepts underlying internal validity and reliability have pendants in credibility (cf. internal validity) and dependability (cf. reliability) as used in qualitative research.¹³ To assess the trustworthiness of qualitative data, Lincoln and Guba have systematically replaced traditional criteria by a set of parallel methodological criteria. A central criterion is credibility, which relates to the truth value within the findings so that they are both believable and supported by the evidence provided.¹³ A number of methodological strategies has been suggested to ensure credibility and dependability.¹⁴ The following 3 strategies are important for reaching credibility: *triangulation* (combining different information sources); *prolonged engagement* (sufficient time investment by the researcher) and *member checking* (testing the data with the members of the group from which they were collected). The strategies for realising dependability – the pendant of reliability- involve establishing an *audit trail* (i.e. documentation of the assessment process to enable external checks) and carrying out a

dependability audit (i.e. quality assessment procedures with an external auditor).

These strategies can also be used to achieve credibility and dependability in educational assessment.¹⁵ For example, Norcini and Shea use the concept of credibility as a criterion against which methods of standard setting can be evaluated.¹⁶ In this approach standard setters, should have the proper qualifications (*prolonged engagement*), many standard setters should be involved and the judgements of other credible groups should be included (*triangulation*).¹³ Credibility strategies are especially useful for constructing an assessment procedure, while dependability strategies can be used to monitor the assessment procedure. We conducted a case study among 237 Year 1 medical students to explore how the concepts of credibility and dependability can be applied to portfolio assessment. We addressed the question as to how such qualitative research strategies can be used in a portfolio assessment procedure to ensure reliable and valid judgement.

Context of the study

This case study explored the portfolio assessment procedure that is used in Year 1 of the undergraduate medical curriculum at Maastricht University, the Netherlands. The structure of the portfolio was provided by 4 different roles of a doctor: medical expert, scientist, healthcare worker and person. Global criteria were devised for each role and students had to collect evidence demonstrating that by the end of the year they had met those criteria. The students were mentored by medical school staff. At the beginning of the academic year the portfolio system was introduced and students carried out some portfolio exercises. In the portfolio students had to present an analysis of their personal strengths and weaknesses in relation to the 4 roles of a doctor. These reflections had to be backed up by evidence, such as feedback from evaluations and tests and completed assignments. Students were also required to draw up a learning plan for the next period. Halfway through the academic year the students submitted their portfolios to their mentors, who gave feedback. In a progress meeting student and mentor discussed the portfolio and the student's competence development regarding the 4 roles. It was assumed that the student would adjust the portfolio in accordance with the feedback received. At the end of the academic year this cycle of submission, feedback and adjustment was repeated. This portfolio format has been described in greater detail elsewhere.¹⁷

Formative and summative assessment

The purpose of the Year 1 portfolio was primarily *formative*. It was intended to promote feedback as part of the assessment programme and help students monitor their competence development and develop reflective, planning and remediation skills.

The portfolio also served a *summative* purpose. This was considered desirable for 2 reasons:

1. because experience has taught that purely *formative* assessment tends to lose momentum and after some time a new impetus is needed to steer student learning into the desired direction,¹⁸ and
2. because portfolio assessment offers a unique opportunity to identify students who are lagging behind in professional progress and who show insufficient ability to reflect, plan and/or take remedial action. As *summative* assessment implies the possibility of failing and students who fail the portfolio may ultimately face expulsion from medical school, it will be clear that the portfolio is a high stakes assessment and fair decisions and maximum prevention of decision errors are of the essence.

In summary, the portfolio assessment procedure in the case study was designed to strike a delicate balance between *formative* and *summative* evaluation, seeking the best possible mix of benefits from both approaches.¹⁹ We will describe how we combined *formative* and *summative* assessment in a single procedure.

The portfolio assessment procedure

All portfolios (n=237) were judged at the end of the academic year and graded as fail, pass or distinction. The following, rather global, criteria were used to assess the quality of the portfolios:

- the quality of the analyses of strengths and weaknesses;
- the quality of the evidence;
- the extent to which the evidence reflected the analyses of strengths and weaknesses;
- the clarity and feasibility of the learning objectives, and
- the extent to which the learning objectives were achieved.

These criteria express the following steps in the portfolio cycle: reflect on competence development; sample evidence; link evidence to reflection; formulate learning objectives, and develop competence. Obviously, such

broad criteria necessitate extensive input from the judges in the assessment process.

Assessment occurred in all phases of the portfolio process:

1. during the compilation of the portfolio in regular meetings of mentor and student;
2. in the end-of-year meeting when mentor and student recommended the final grade, and
3. after submission of the portfolio to the Portfolio Assessment Committee (PAC) for final grading.

In all 3 phases procedures and precautions had to be in place to ensure a credible assessment process.

Compiling the portfolio

Over the course of the year the students discussed their progress as documented in the portfolio in at least 2 sessions with their mentor, who provided constructive oral and written narrative feedback.

The mentor's combined role of supervisor and assessor can be a difficult, albeit not impossible, task. A classic example of a situation involving a similar role combination is the relationship between supervisor and PhD student. The supervisor has to coach and encourage the student to put his/her best efforts into the dissertation, while at the same time both supervisor and student have to prevent submission of an unsatisfactory dissertation to an external assessment panel. In our case study, this analogy has been particularly useful in the training sessions for the mentors at the beginning of the academic year. These sessions were supplemented by intervision sessions, in which the mentors shared experiences and information. The purpose of this approach was to support the mentors in their difficult double role and build a sound foundation for the feedback to the students. Another advantage of regular feedback is that it prevents students being disappointed by an unexpected, negative recommendation to the portfolio committee from their mentor.

Recommendation by mentor and student

In their final meeting of the academic year the student and the mentor discussed the mentor's well-motivated recommendation to the assessment committee concerning the grading of the portfolio. When student and mentor agreed on the grade, the student signed the recommendation. The student did not sign if there was disagreement, which the student indicated on the

assessment form. Subsequently, the portfolio was submitted to the committee.

Portfolio Assessment Committee

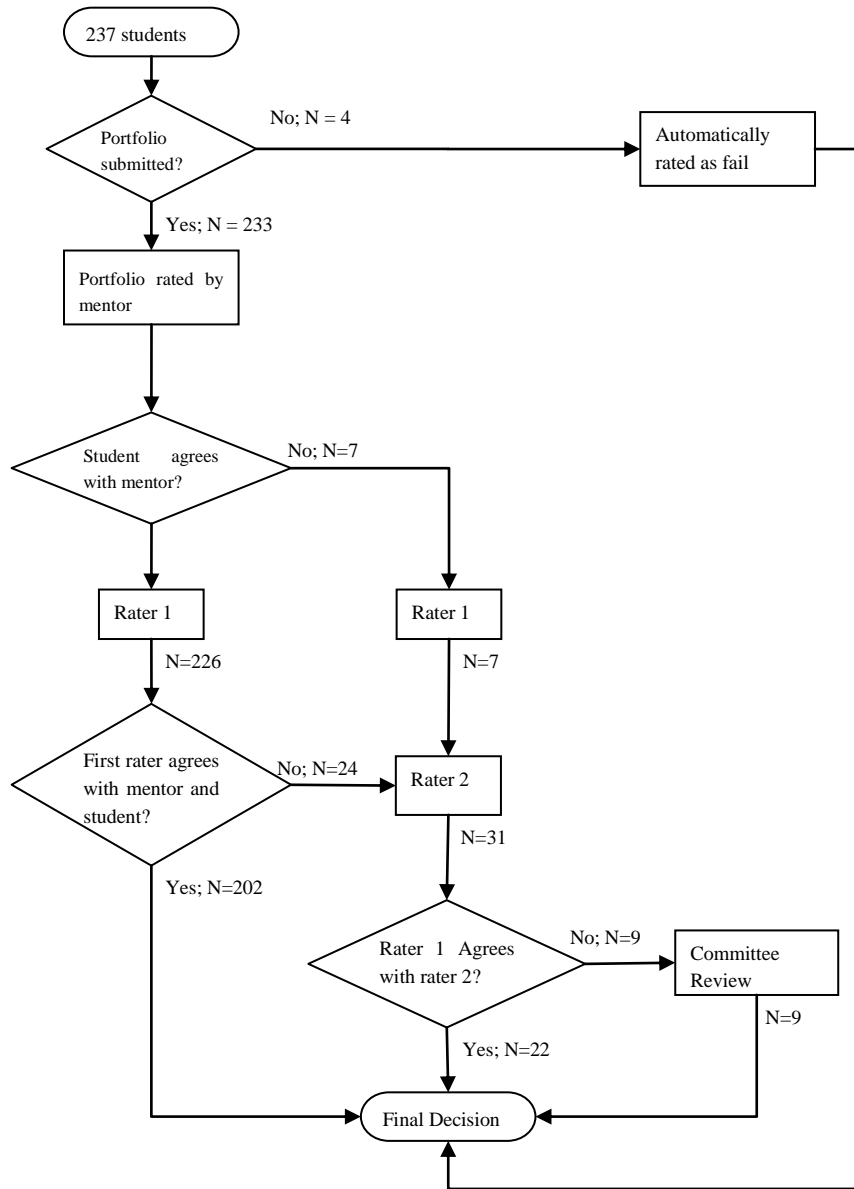
The final step of the assessment procedure comprised a sequential judgement procedure by the assessment committee. As it is the mentors who have first-hand experience with the portfolios, it was decided that the assessment committee should be composed of the 13 Year 1 mentors. The committee members did not grade the portfolios of the students they had mentored. Because judging a portfolio is time-consuming, the assessment procedure was designed for maximum efficiency. Judgements of the full committee were only required if the available information was not unanimous. Figure 1 presents the assessment procedure in a flowchart.

The flowchart shows the number of portfolios remaining to be judged after each step in the decision process. A total of 233 portfolios were submitted to the assessment committee at the end of the academic year 2001-02. Firstly, the portfolios on which student and mentor agreed were rated by a single committee member, who did not study the portfolio in any great detail, but typically scanned the work of the student and mentor and checked whether all procedures had been followed correctly. Only if the rater had any doubts, was the portfolio examined further. When rater, mentor and student agreed on the grading, the recommendation became the final decision. The fact that this parsimonious route proved feasible for the majority of portfolios ($n=202$; 85%), may be taken as an indication that the assessment procedure, with mentor-student meetings, was satisfactory. If the rater did not agree with student and mentor, a second independent rater judged the portfolio. If the 2 raters agreed, their judgement became the final decision. If they disagreed, the portfolio was submitted to the full committee.

Students and mentors disagreed about the grading of only 7 (3%) portfolios. These portfolios were judged by 2 raters independently of one another. When the 2 raters agreed, their rating became the final decision. If they disagreed, the portfolio was submitted to the full committee.

Only 9 portfolios (4%) were discussed in the full committee meeting. These portfolios were reviewed individually, with mentor and raters presenting their arguments. The final decision was based on consensus among the committee members, excluding the student's mentor. As the mentor is first and foremost supervisor, rather than an assessor, the mentor had no vote in the final decision. In all, 29 (12%) final decisions differed from the original recommendation. Nine students failed, 147 received a pass and 81 were given a distinction.

Figure 1: Flowchart of the judgement procedure of the portfolio assessment committee.



A total of 226 portfolios (96%) were graded without being reviewed by the full committee. The entire procedure was completed in the relatively short time of 42 hours (i.e. 11 minutes per portfolio, with the committee meeting lasting 1 hour). The participants did not perceive the process as particularly stressful.

Discussion

This case study demonstrates the feasibility of a qualitative approach to achieve reliable summative judgement using an inherently complex and non-standardised assessment instrument, which relies on holistic professional judgement. We incorporated some procedural safeguards into the assessment process to achieve maximum credibility of the decisions.¹³⁻¹⁵ Essential elements in the assessment process were:

- feedback cycles, incorporated into the mentoring process during the compilation of the portfolio to ensure that the mentor's final recommendation did not come as a(n) - unpleasant - surprise to the student; this element relates to the credibility strategies of prolonged engagement and member checking;
- maintaining a careful balance between the mentor's roles of coach and assessor, ensuring that the person who knew the student best provided the most relevant information while at the same time minimising any damaging effect to the mentor-student relationship; this relates to the credibility strategy of prolonged engagement;
- student involvement in the decision process to ensure commitment on the part of the student and allow the student to communicate a different point of view to that of the mentor; this relates to the credibility strategy of member checking, and
- a sequential judgement procedure in which conflicting information necessitated more information gathering, ensuring efficient use of resources by reserving efforts to achieve more reliable judgement in cases where this was absolutely necessary. As a result, more resources (i.e. mentor time) were available for coaching students, which is in line with the main purpose of the portfolio. This element relates to the credibility strategy of triangulation.

Dependability can be reached by establishing an audit trail and by the use of external auditors. Both strategies were used to monitor our assessment procedure. The audit trail consisted of comprehensive documentation of the different steps of the assessment process: a formal assessment plan approved

by the examination board; portfolio and assessment guidelines; overviews of the results per phase, and written assessment forms per student. Quality assurance procedures were set up. The internal quality procedure enables a student to appeal to the University Board of Appeal for the Examinations against the outcome of the assessment. The external quality procedure entails regular audit by the Dutch organisation for educational auditing and accreditation. This relates to the dependability strategy of audit.

The case study showed that all these elements contributed to the credibility and dependability of portfolio assessment. We are convinced that this assessment process has considerably more credibility and dependability than procedures aimed at high inter-rater reliability, particularly if such procedures necessitate standardisation and rigid structuring with concomitant impairment of validity. In the hypothetical case of a legal procedure, we would be able to build our defence on the evidence that after a careful assessment procedure a committee of experts had reached consensus on the final decision on a student's portfolio. A better defence is hard to imagine. Further support for this assessment procedure might be sought by means of a study involving different assessment committees.

Although the mentoring process was resource intensive most of the mentors' time was spent on mentoring and formative feedback and only a minor portion on formal assessment. During an informal debriefing the mentors indicated that the judgement procedure had not burdened them disproportionately.

To some extent, the portfolio assessment procedure described in this paper resembles the procedure suggested by Friedman Ben David *et al.*, who proposed 2 independent ratings followed by a final consensus procedure, culminating in an overall judgement and agreement between the 2 raters.¹ Our procedure is broader and includes information collected from the very start of the portfolio compilation process, the mentor's and the student's input as well as a final full committee consensus decision.

At the heart of the approach used in our case study is the recommendation by McCullan *et al.* in their review of portfolio studies, that, for portfolio assessment, criteria from qualitative research might be more appropriate than criteria from quantitative research, like reliability.³ Instead of looking at consistency across repeated assessments (a quantitative psychometric approach), we added information to the judgement process until saturation of information was reached (a qualitative approach).²⁰ This does not mean that

psychometric aspects were ignored. In fact, the concept of psychometrics – particularly in relation to sequentially increasing the number of examiners – was used, but it was not applied in a classic test theoretical sense.

Naturally, some arbitrary decisions were unavoidable. It is not unlikely that adaptations of the procedure will be proposed, depending on our cumulative experiences and evaluations. However, the essence of our argument is that we have tried to demonstrate that reliability, as viewed from the purely psychometric perspective, is too limited a criterion to be applied to qualitative assessment and that reliability can be built into an assessment procedure by implementing various safeguards and by collecting more information only when necessary. We believe that this represents an important step forward in our endeavours to incorporate more qualitative and subjective features into competence assessment.

References

1. Friedman Ben David M, Davis MH, Harden RM, Howie PW, Ker J, Pippard MJ. AMEE Medical Education Guide No. 24: Portfolios as a method of student assessment. *Med Teacher* 2001;23(6):535-51.
2. Webb C, Gray M, Jasper M, McCullan M, Scholes J. Models of portfolios. *Med Educ* 2002;36:879-98.
3. McCullan M, Endacott R, Gray MA, Jasper M, Miller CML, Scholes J, et al. Portfolios and assessment of competence: a review of the literature. *J Adv Nurs* 2003;41(3):283-94.
4. Snadden D. Portfolios - attempting to measure the unmeasurable? *Med Educ* 1999;33:478-9.
5. Baume D, Yorke M. The reliability of assessment by portfolio on a course to develop and accredit teachers in higher education. *S Higher Educ* 2002;27(1):7-25.
6. Pitts J, Coles C, Thomas P. Educational portfolios in the assessment of general practice trainers: reliability of assessors. *Med Educ* 1999;33:515-20.
7. Roberts C, Newble DI, O'Rourke A. Portfolio-based assessments in medical education: are they valid and reliable for summative purposes? *Med Educ* 2002;36:899-900.
8. Koretz D, Stecher B, Klein S, McCaffrey D. The Vermont Portfolio Assessment Program: findings and implications. *Educational Measurement: Issues and Practice* 1994;13(3):5-16.
9. Van der Vleuten CPM, Van Luijk SJ, Ballegooijen AMJ, Swanson DB. Training and experience of medical examiners. *Med Educ* 1989;22:290-6.
10. Norman GR, Van der Vleuten CPM, De Graaff E. Pitfalls in the pursuit of objectivity: Issues of validity, efficiency and acceptability. *Med Educ* 1991;25:119-26.
11. Huot B. Beyond the classroom: using portfolios to assess writing. In: Black L, Daiker DA, Sommers J, Stygall G, editors. *New directions in portfolio assessment: Reflection, practice, critical theory and large-scale scoring*. Portsmouth, NH: Cook Publishers; 1994. p. 325-33.
12. Van der Vleuten CPM, Norman GR, De Graaff E. Pitfalls in the pursuit of objectivity: Issues of reliability. *Med Educ* 1991;25:110-8.
13. Lincoln YS, Guba EA, editors. *Naturalistic inquiry*. Beverly Hills, CA: Sage; 1985.
14. Denzin NK, Lincoln YS, editors. *Handbook of Qualitative Research*. 2nd ed. Thousand Oaks, CA: Sage; 2000.

15. Webb CER, Gray M, Jasper M, McMullan M, Scholes J. Evaluating portfolio assessment systems: what are the appropriate criteria? *Nurse Educ Today* 2003;23:600-9.
16. Norcini JJ, Shea JA. The credibility and comparability of standards. *Appl Measurement Educ* 1997;10(1):39-59.
17. Driessen EW, Van Tartwijk J, Vermunt JD, Van der Vleuten CPM. Use of portfolios in early undergraduate medical training. *Med Teacher* 2003;25(1):14-19.
18. Driessen EW, Van der Vleuten CPM. Matching student assessment to problem-based learning: lessons from experience in a law faculty. *Studies Cont Educ* 2000;22(2):235-48.
19. Zeichner K, Wray S. The teaching portfolio in U.S. Teacher education programs: What we know and what we need to know. *Teach Teach Educ* 2001;17(5):613-21.
20. Strauss AL. *Qualitative analysis for social scientists*. New York: Cambridge University Press; 1987.

4 Conditions for successful reflective use of portfolios in undergraduate medical education

**Erik W Driessen
Jan van Tartwijk
Karlijn Overeem
Jan D Vermunt
Cees PM van der Vleuten**

Published in: Medical Education 2005; 39: 1230-1235

Summary

Aim

Portfolios are often used as an instrument to stimulate students to reflect on their experiences. Research has shown that working with portfolios does not automatically stimulate reflection. In this study we addressed the question: “what are the conditions for successful reflective use of portfolios in undergraduate medical education?”

Methodology/research design

We designed a portfolio that was aimed at stimulating reflection in early undergraduate medical education, using experiences described in the medical education literature and elsewhere. Conditions for reflective portfolio use were identified through interviews with 13 teachers (mentors), who were experienced in mentoring students in the process of developing their portfolios. The interviews were analysed according to the principles of grounded theory.

Results

The conditions for successful reflective use of portfolios that emerged from the interviews fell into 4 categories: coaching; portfolio structure and guidelines; relevant experiences and materials, and summative assessment. According to the mentors, working with a portfolio designed to meet these conditions will stimulate students’ reflective abilities.

Conclusion

This study shows that portfolios are a potentially valuable method of assessing *and* developing students’ reflective skills in undergraduate medical training, provided certain conditions for effective portfolios are recognised and met. Portfolios have a strong potential for enhancing learning and assessment but they are very vulnerable and may easily lead to disappointment. Before implementing portfolios in education, one should first consider whether the necessary conditions can be fulfilled, including an appropriate portfolio structure, an appropriate assessment procedure, provision of enough new experiences and materials, and sufficient teacher capacity for adequate coaching and assessment.

Introduction

Portfolios are widely used in healthcare education, not only as a source of information for authentic assessment but also to help students reflect on their experiences.¹ We regard reflection as a cyclic process of self-regulation in which students look back on their actions, analyse them, think up alternatives, try these out in practice, look back on them, etc.² The objective of this process is to learn from experience. Reflection thus becomes a condition for professional development.²

Research has shown that a portfolio is no guarantee that reflection will occur. Pearson and Heywood investigated the reflective use of a portfolio by general practice registrars.³ This portfolio was “an A5 document containing pages for use as a learning record and including prompts for the identification of learning needs together with space to record teaching and assessment information”.³ Pearson and Heywood considered as evidence of reflection “... recording information ... combined with active recollection”, consisting of “... either a memory of self-reflection on the information, or discussion of the information with the trainer”.³ It appeared that the majority of the registrars and their trainers did not use the portfolio for reflection and that its purpose was unclear. Moreover, many trainers were not motivated to work with the portfolio and thus did not encourage its use. This correlated with poor use of the portfolio for reflection. Finally, the strictly prescribed format further hampered acceptability.

Research in teacher education has revealed that portfolios only stimulate reflection if certain conditions are met. According to Wade and Yarbrough, reflective portfolio use stimulates students to develop new understandings and appreciations of their experiences, recognise links between different aspects of these experiences and formulate insights to be tested in future actions.⁴ The portfolio in their study consisted of a set of artefacts representing a range of decisions, a record of each student’s personal development and a reflective essay. Wade and Yarbrough identified the following conditions for reflective portfolio use: a good introduction of the portfolio and its intended use, student ownership, a clear structure, and appropriate use of the portfolio in discussions with coaches or trainers.⁴

Reflection is a prerequisite for learning in the context of real practice, which is a prominent feature of many innovative medical curricula. A portfolio can

foster reflective skills, provided the learning environment is favourable. So the question is how to create such an environment. To explore this we addressed the following research question: what are the conditions for successful use of portfolios for reflection in undergraduate medical education? We designed a portfolio intended to stimulate reflection in early undergraduate medical education. The design was based on experiences with portfolios described in the medical education literature and elsewhere.⁵ We sought teachers' views on determinants of effective portfolios by interviewing teachers (mentors) who had had considerable experience in discussing portfolios with students. We focused on experienced teachers' views, not only because we believe that they have the most accurate insights into the benefits and pitfalls of reflective portfolio use, but also because teachers' perceptions of the use and usefulness of portfolios appear to be a decisive factor in successful implementation.^{3,4}

Methods

Context of the study

In the curriculum of Maastricht Medical School, the Netherlands, authentic learning is introduced in Year 1, where students' learning environment consists of authentic and real cases. Although portfolios are predominantly used at more senior levels of training, it was decided that it would be appropriate to introduce a portfolio early in the curriculum to help students develop reflective ability. The portfolio consisted of 3 parts:

- written self-assessments of students' personal development in 45 professional roles and the learning goals students derive from these self-assessments to further their role development;
- illustrative materials underpinning the self-assessments, and
- (written) feedback by students' personal mentors on the first 2 parts.

Every year students discuss their portfolios in at least 2 one-on-one meetings with their personal mentors. Mentors evaluate portfolio quality and address aspects that need improvement. Their written feedback is added to the portfolios.

Students' reflective skills are assessed (pass/fail) annually by the Portfolio Committee. In the preclinical phase, reflective ability is assessed. In the clinical phase, portfolios contribute to the assessment of students' clinical performance. All mentors sit on the Portfolio Committee, but mentors do not assess the portfolios of the students they are mentoring.

A detailed description of the portfolio and the assessment procedure can be found elsewhere.^{5,6}

Mentors

We interviewed all mentors (n=13) who were coaching Year 1 and 2 students during the period of the study. Over a 2-year period, they mentored an average of 20 undergraduate students using the portfolio described above. All mentors had more than 5 years of experience in teaching small groups and individual students. Four of the mentors were women. Participation in the study was voluntary and participants received a small fee and a short report of the finished study.

Interviews

The interviews lasted approximately 1 hour and were tape-recorded after consent had been obtained from participants. Literal transcriptions were made.

Three topics were addressed in each interview:

- the mentor's (implicit) definition of reflective skills;
- the portfolio's effectiveness in stimulating students to reflect on their experiences and development, and
- conditions for successful reflective portfolio use.

Analyses

The interviews were analysed according to the principles of grounded theory using the program ATLAS/ti. After the first 5 interviews, 3 interviews were coded independently by 2 researchers (EwD and JvT). Coding comprised selecting citations and assigning labels to them. The outcomes were compared and any differences discussed until consensus was reached. The researchers then re-read the interviews to check that no relevant information had been overlooked. The resulting refined interview schema was used in the interviews with the remaining 8 mentors. The final step of the analysis was member checking (i.e. determining whether interviewees agreed that data and conclusions accurately reflected interview content).⁷ For this purpose, 2 of the interviewees read and commented on the results and conclusions of the study. This part of the analysis did not necessitate any changes.

Results

We report the results for the main interview topics, which were definition of reflective skills, the effectiveness of the portfolio in stimulating these skills and conditions for successful portfolio use.

Definition of reflective skills

In the eyes of most mentors, reflection focuses on professional attributes and its purpose is to offer directions for improvement by identifying strengths and weaknesses in performance. The mentors saw reflection as a method of identifying causes in order to answer the question of why things are as they are. The “why-question” was regarded as essential. One mentor described a portfolio without reflection:

“...as a collection of facts without questioning the whys and wherefores.”
(mentor 1)

Mentors attached great importance to students addressing why-questions from a position outside their personal perspectives. They saw openness to the opinions of others as a prerequisite for objectivity.

“Normally, you observe the environment and yourself from a personal perspective. When reflecting, you look at yourself from the perspectives of others. This means for example appreciating others’ opinions and trying to understand them.” (mentor 11).

“That you can objectify your own attitude and behaviour and try to eliminate all bias so as to obtain a clear view of your position and understand how you are seen by others.” (mentor 2)

Some mentors wondered how far they should go in asking students to self-reflect. As one mentor said:

“Self-reflection may cause students to recognise or become aware of situations in their personal lives that are less than ideal. The question is how to deal with this. The question is also whether this is desirable and whether it is appropriate.” (mentor 2)

Effectiveness of the portfolio in stimulating students to reflect on their experiences and development

All mentors said that compiling portfolios and writing reflective reports fostered a critical attitude in students towards their own performance and helped them manage their own development. The mentors said that the

portfolio not only stimulated students to examine their experiences retrospectively and systematically, it also offered directions for development. This process was seen as affecting the pace of students' development:

“Some of the things I myself had to learn the hard way could have been dealt with at a much earlier stage of my medical training with the use of a portfolio system.” (mentor 4)

The mentors said that ability, attitude and motivation determined how easily students learned to reflect.

“I think there are 3 types of students who are poor portfolio compilers. You have the students who lack analytical ability. You have the group of students who think compiling a portfolio is completely pointless; the characters of these 2 student groups are incompatible with reflection and portfolios. Finally, you have the students who just aren't motivated.” (mentor 13)

The need to identify those students who lack the ability to critically appraise their own performance was emphasised by the mentors, because this is considered to be a vital skill for medical doctors. Discussing portfolios with students is 1 way of identifying this group.

Conditions for successful reflective portfolio use

Four categories of conditions for successful reflective portfolio use emerged from the interviews: good coaching; structure and guidelines; adequate experiences and material for reflection, and summative assessment.

Coaching

Coaching plays a crucial role in reflection. As reflection does not come naturally to most students, it is an important task for coaches or mentors to show students what questions to ask themselves when reflecting on their performance. Several mentors said they had not realised this was difficult for many students. Mentor 12 put it this way:

“I didn't know self-analysis was so difficult for students. I thought that everybody would include a self-analysis in their portfolios.” (mentor 12)

Another important function of coaching involved helping students to identify learning needs and design learning plans. According to the mentors, these learning plans were often no more than a list of resolutions, rather than well thought out, realistic steps towards relevant learning goals.

Student motivation was also mentioned as an important factor in successful reflective portfolio use. The mentors indicated they sometimes had to go to considerable lengths to convince students of the benefits of reflective portfolio use:

“I explain to them what my idea of reflection is and I try to make them see the value of the portfolio and sometimes I show them something from another student’s portfolio.” (mentor 1)

Structure and guidelines

According to the mentors, a portfolio should be well-structured and guidelines should tell students what is expected of them and what are suitable subjects for reflection. Several mentors also pointed out that a more open portfolio structure was preferable once students had learned how to reflect. Although weaker students needed structure and guidelines, too much structure may become an obstacle for students with good reflective skills. These students should have more freedom in compiling their portfolios.

Experiences and material

Another condition for an effective portfolio was sufficient variety and quantity of interesting experiences as subjects for reflection. The mentors reported that lack of experiences was problematic if student-mentor meetings were planned too closely together. If students have no subjects for reflection, it becomes a rather futile exercise:

“It seems to me that they are actually repeating themselves, i.e. they are doing what they have to do just because it is obligatory”. (mentor 8)

Summative assessment

The mentors indicated that the portfolio should be used for summative assessment of reflective skills to ensure that it is taken seriously by students and mentors. Reflective portfolio use is labour intensive for both students and mentors. There was agreement among the mentors that, without assessment, portfolios would be taken less seriously and students and mentors would not consider it worthwhile to invest the necessary time and energy.

“I really think students take the portfolio much more seriously when the portfolio is assessed.” (mentor 5)

“It prevents you from thinking: Oh well, I’ll have a talk with this student and that’s my main goal and I don’t need a portfolio. It stops you from failing to take the portfolio seriously, I think.” (mentor 13)

Some mentors argued that test-directed portfolio compilation was a potential pitfall of summative assessment. This is a very real danger if the above-mentioned conditions are not met, if there are not enough new and relevant experiences for students to reflect on, if students are given too many rules and guidelines, or if portfolio content is not the subject of serious discussions with mentors. Mentors generally thought that students were “honest” in their portfolios and “open” in student-mentor meetings. They did not think students looked upon them as assessors.

“The fact that the portfolio is assessed has in my opinion had no effect whatsoever on the openness from both sides.” (mentor 9)

Discussion

The interviewees agreed that compiling and discussing portfolios enhanced the development of students’ reflective abilities. The findings suggest that, although reflection may initially be difficult for Year 1 students, almost all students can learn how to reflect, provided favourable conditions are created. Only a small minority fails to develop reflective skills. According to the mentors, it is important to identify this minority, because the ability to critically appraise one’s performance is a crucial professional skill for medical doctors.

A limitation of this study is that we examined the teachers’ perspective only and not that of the students.

The mentors thought the success of portfolio learning depended on 4 categories of conditions: good coaching; appropriate portfolio structure and guidelines; a sufficient number of relevant experiences on which to reflect, and summative assessment. The first 2 categories have also been reported in other studies.^{3,4}

A supportive teacher (mentoring) or peer feedback is widely recognised as a key factor in the success of reflective portfolios.^{3,4,8} This was confirmed by the results of our study. Pearson and Heywood even went so far as to say that reflection was primarily stimulated by supervisors and that a portfolio offered little added value as a stimulus for reflection.³

The same authors reported that the purpose of the portfolio they studied was rather unclear, trainers were poorly motivated and the portfolio format was too rigid.³ Similarly, Wade and Yarbrough identified structure as an

important factor in the effectiveness of a reflective portfolio.⁴ In our study, mentors advocated a careful balance between a highly prescribed structure and full freedom. A well-defined structure was thought to be mainly helpful at first, when lack of structure might cause frustration, especially when students are new at compiling a portfolio.⁹ However, students with good reflective skills should be offered leeway to display their personal qualities. As Wade and Yarbrough put it, students should have ownership of their portfolios.⁴

The mentors also pointed out that the effectiveness of the portfolio depended on students having different and meaningful experiences on which to reflect. This may well be a decisive factor in the successful introduction of a portfolio in early undergraduate medical education. In the first pre-clinical years, most students have fewer relevant and salient experiences than during clinical training. Students who are told to reflect without being exposed to challenging experiences can easily become "*portfolio tired*", a phenomenon also encountered by Snadden and Thomas in the context of General Practice training.⁸

The results of our study show that summative assessment prevents students and mentors from adopting too casual an attitude towards the portfolio. Effective portfolios require substantial effort and time from students and mentors. If a portfolio is not graded, students and mentors may question whether the whole exercise is worth the effort. In the literature we found some examples of disappointing portfolio experiences that may be attributable to this phenomenon.^{8,10} Snadden and Thomas reported a study in which a portfolio of general practice trainees was not assessed summatively.⁸ Trainees appeared to stop working on their portfolios as soon as other, summative, assessments made demands on their time.

Reflection was characterised by the mentors as thought processes that help students improve their professional performance. The prevailing opinion was that reflection should focus on actions - and their consequences - which are directly related to students' future work as medical professionals. In other fields, such as teacher education, more emphasis tends to be placed on student-teachers' individual identities and beliefs, because these are regarded as crucial to their professional development.¹¹ The mentors we interviewed did not focus on students' personal opinions and motives in portfolio discussions.

The results of our study suggest that portfolios are a potentially valuable method for the assessment and development of undergraduate medical students' reflective skills. Portfolios can be a powerful tool for learning and assessment, but the method is vulnerable to adverse conditions and may easily lead to disappointment. Those intending to implement a portfolio should carefully consider whether they will be able to create the favourable learning environment needed for successful portfolio use.

References

- 1 Rees C, Sheard C. Undergraduate medical students' views about a reflective portfolio assessment of their communication skills learning. *Med Educ* 2004;38:125-8.
- 2 Korthagen FAM. *Linking practice and theory*. Mahwah, N.J.: Lawrence Erlbaum Associates Mahwah, N.J. 2001.
- 3 Pearson DJ, Heywood P. Portfolio use in general practice vocational training: a survey of GP registrars. *Med Educ* 2004;38:87-95.
- 4 Wade RC, Yarbrough DB. Portfolios: a tool for reflective thinking in teacher education? *Teaching & Teacher Education* 1996;12(1):63-79.
- 5 Driessen EW, Van Tartwijk J, Vermunt JD, Van der Vleuten CPM. Use of portfolios in early undergraduate medical training. *Med Teach* 2003;25(1):14-19.
- 6 Driessen E, Van der Vleuten C, Schuwirth L, Van Tartwijk J, Vermunt J. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. *Med Educ* 2005;39(2): 214-220.
- 7 Lincoln YS, Guba, G. *Naturalistic inquiry*. Newbury Park: Sage; 1985.
- 8 Snadden D, Thomas ML. Portfolio learning in general practice vocational training - does it work? *Med Educ* 1998;32:401-6.
- 9 Gearhart MG, Wolf SA. (1997) Issues in portfolio assessment: assessing writing processes from their products. *Educational Assessment* 1997;4(4):265-96.
- 10 Finlay IG, Maughan TS, Webster DJT. A randomized controlled study of portfolio learning in undergraduate cancer education. *Med Educ* 1998;32:172-6.
- 11 Korthagen F.A.J. (2003). In search of the essence of a good teacher: toward a more holistic approach in teacher education. *Teaching and Teacher Education* (20) 77-97.

5 Validity of portfolio assessment: which qualities determine ratings?

**Erik W Driessen
Karlijn Overeem
Jan van Tartwijk
Cees PM van der Vleuten
Arno MM Muijtjens**

Published in: Medical Education 2006; 40: 862-866

Summary

Introduction

Portfolios are increasingly being accepted as a valuable tool for learning and assessment. The validity of portfolio assessment, however, may suffer from bias due to irrelevant qualities, such as lay-out and writing. We examined possible effects of such qualities in a portfolio aimed at stimulating Year 1 medical students to reflect on their professional and personal development. In later curricular years, this portfolio is also used to judge clinical competence.

Methods

We developed an instrument, the Portfolio Analysis Scoring Inventory, to examine the impact of form and content aspects on portfolio assessment. The Inventory consists of 15 items derived from interviews with experienced mentors, the literature, and the criteria for reflective competence used in the regular portfolio assessment procedure. Forty portfolios, selected from 231 portfolios for which ratings from the regular assessment procedure were available, were rated by 2 researchers, independently, using the Inventory. Regression analysis was used to estimate the correlation between the ratings from the regular assessment and those of the Inventory items.

Results

Inter-rater agreement ranged from 0.46 to 0.87. The strongest predictor of the variance in the regular ratings was 'quality of reflection' (R 0.80; R² 64%). No further items accounted for a significant proportion of variance. Irrelevant items, such as writing style and lay-out, had negligible effects.

Conclusions

The absence of an impact of irrelevant criteria appears to support the validity of the portfolio assessment procedure. Further studies should examine the portfolio's validity for the assessment of clinical competence.

Introduction

It is the strength of portfolios that they offer rich and authentic evidence of learners' development and achievements. This makes them highly suitable not only for monitoring, but also for assessing learners' competence development. High validity is generally attributed to portfolio assessment.¹ Portfolios afford insight into learners' clinical competence through authentic evidential materials collected by learners in day-to-day practice over a prolonged period of time. Unfortunately, published studies on validity issues in relation to portfolios are rare. Indeed, the validity of portfolio assessment has been much less studied than the reliability of this type of assessment, particularly in medical education.² The few studies that we found in the literature lend some support to different aspects of portfolio validity: criterion and construct validity, predictive validity, and content validity.³⁻⁵ Given that sound decisions require assessment that is both reliable and valid, the validity of portfolio assessment warrants further study.

Portfolio assessment is complex. This complexity is inherent to the open format of portfolios. An assessor has to judge portfolios that differ in content, size and, in many cases, structure. The richness and complexity of portfolios cannot be captured by analytic assessment criteria and detailed checklists can easily trivialise assessment.⁶ That is why global (holistic) assessment methods characterised by strong reliance on assessors' professional judgements have been advocated for portfolio assessment.^{7,8} A potential source of bias in such assessment procedures is that assessors may be tempted to let irrelevant qualities, such as the quality of the writing, the structure and the lay-out of the portfolio, sway their judgement. Presentation and students' personal characteristics have been shown to interfere with portfolio assessment, because they may mistakenly be interpreted as signs of competence in the area of interest.¹ For instance, in a study on portfolio assessment in teacher education by Quinlan, the analysis of thinking-aloud protocols revealed that competence ratings were influenced by what assessors already knew about students.⁹ Heller *et al.*, who also used thinking aloud protocols, stated: "what is essential for maintaining valid scores is that raters be capable of consciously discriminating between relevant and irrelevant qualities during the rating judgement".¹⁰

Because of the real risk that bias due to irrelevant portfolio qualities might compromise the validity of portfolio assessment, we designed a study to

examine which criteria determined portfolio assessment of reflective competence. Our research questions were:

- Which criteria affect raters' judgments of students' reflective skills?
- Which criteria carry the most weight?

Methods

Context

At Maastricht University, the Netherlands, portfolios are currently used to stimulate the development of students' academic and professional competencies over the course of the 6-year undergraduate medical curriculum.¹¹ Early in the curriculum, the portfolio's prime focus is on reflective skills, because reflection is regarded as a prerequisite for effective learning from experience. In this context, reflection is seen as a cyclic process. It starts with students analysing their learning experiences. In the next step, students distil learning objectives from the strengths and weaknesses emerging from the analysis. These learning objectives are then pursued by the students during subsequent experiences, at which point the cycle of analysing experience and generating and pursuing learning objectives starts afresh, to be repeated again and again. It is the role of the portfolio to invite students to record and therefore put into words what they perceive as the strengths and weaknesses in their performance as well as their learning objectives and whether and how they have attained those. In addition, students meet with their personal mentors at least twice a year to talk about the portfolio. The mentors try to steer students' efforts into fruitful directions. Instructions and guiding questions are another way to enhance students' reflective processes. The portfolio template offers some degree of structure and uniformity: students are expected to record the outcomes of reflecting on their performance in 4 professional roles: medical expert; health care professional; scholar and person.¹² At the end of the year the portfolios are judged in a summative holistic assessment procedure based on the following (global) criteria⁸:

1. that the student's analyses of strengths and weaknesses of role performance in the different roles are appropriate;
2. that the student has generated learning objectives that are clearly defined and feasible;
3. that the student has attained the learning objectives to a sufficient degree;
4. that the students provides appropriate evidence to support the analyses of strengths and weaknesses, and

5. that the portfolio contains all the required items and was handed in on time.

The mentors are trained in coaching students' portfolio and reflective skills at the start of mentorship. Just before the end-of-year assessment the mentors are trained in assessing the portfolios. The training consists of discussing and benchmarking portfolios from a previous year. The end-of-year assessment yields a rating of reflective competence as poor (fail), satisfactory or good. Mentors propose ratings and discuss these with the students, who are asked to express agreement or disagreement. Subsequently, each portfolio is assessed by an other mentor and the final rating is determined by the Assessment Committee, which is composed of all the mentors. The Assessment Committee discusses only those portfolios that give rise to differences of opinion between mentor, student and/or other mentor.⁸

Instrument

In order to study the effects of 'irrelevant' qualities on portfolio ratings, we developed an instrument, designated the Portfolio Quality Analysis Scoring Inventory. The 15 items of this Scoring Inventory are based on interviews conducted by the first author with the mentors of Year 1 medical students about criteria for portfolio content, structure and presentation, and the above-mentioned assessment criteria for reflective competence.¹² The items consist of propositions to be rated on a Likert scale from 1 = strongly disagree to 5 = strongly agree (Table 1). Scoring instructions explain when the different Likert scores are appropriate.

In a pilot of the Scoring Inventory, the second author and 1 of the mentors rated 10 portfolios using the Inventory. All 15 items were found to be relevant, but inter-rater agreement was unsatisfactory on 8 items. The wording of these items was revised to make them more accurate and precise. The scoring instructions were also refined.

Procedure

We collected a stratified sample of 40 portfolios out of a total of 231 portfolios compiled by Year 1 students. These portfolios had already passed through the end-of-year assessment procedure described above. The sample was reasonably representative as regards the initial ratings and distribution across different mentors (Table 2). The students gave permission for the use of their portfolios for research purposes.

Table 1. Inter-rater agreement for the Inventory items (significance $p < 0.005$; except for item 2 significance $p < 0.025$)

Inventory items	Weighted kappas
1. This portfolio is among the best as regards lay-out.	0.62
2. This portfolio is among the best as regards spelling and sentence structure.	0.46
3. This portfolio is well-structured, i.e. content is presented in the proper place, descriptions, analyses and learning objectives are easy to find.	0.64
4. The portfolio is complete, i.e. no required components are missing.	0.87
5. The student has looked critically at him/herself, i.e. indicates both strengths and weaknesses for the roles on which work was done.	0.72
6. The analyses of strengths and weaknesses include a search for both internal and external explanations. The analysis is not limited to an enumeration of facts and/or situations.	0.71
7. The analyses of strengths and weaknesses contain a sufficient number of different themes for each role.	0.70
8. The student refers to evidence included in the portfolio in a systematic fashion, i.e. the evidence supports the analyses of strengths and weaknesses.	0.63
9. The student has made a connection between extracurricular activities and the development of his/her competencies.	0.77
10. As for portfolio content, the student has done more than merely following the guiding questions.	0.62
11. The student refers to earlier versions of the portfolio (what went wrong, what went well this time and why, which statements did I make earlier).	0.62
12. The student has formulated logical (following from the analyses of strengths and weaknesses) and clearly defined learning objectives.	0.70
13. The student explains how s/he wants to achieve the learning objectives.	0.70
14. The student has attempted to show what s/he has undertaken to achieve the learning objectives.	0.53
15. The student has expended more effort on the portfolio than was absolutely necessary.	0.61

Table 2. Results of the regular portfolio assessment procedure for all Year 1 portfolios and the study sample

Rating	Year 1 (n=231)	Sample from Year 1 (n=40)
Poor	25 (10.82%)	5 (12.50%)
Satisfactory	124 (53.60%)	20 (50.00%)
Good	82 (35.50%)	15 (37.50%)

The second author and a teacher who was otherwise not involved in the study independently rated the 40 portfolios using the Portfolio Quality Analysis Rating Inventory. To enhance inter-rater reliability, the raters discussed the items after rating 2 portfolios. These portfolios were not included in the study. The raters were blinded to results of the regular assessment procedure.

Data analysis

We used the mean Inventory score across assessors as an indicator of the quality of the portfolio. The reliability of this indicator was obtained by calculating the inter-rater agreement using weighted Kappa statistics (using “absolute error weights”).^{13,14}

We used stepwise multiple regression analysis to determine which of the Inventory items best predicted the ratings of reflective competence.

Results

Inter-rater agreement was acceptable ranging from 0.46 (moderate agreement) to 0.87 (excellent agreement) (Table 1).¹⁵

The standardised regression coefficients (beta) reflect the relative contributions of the independent variables (Table 3). Significance was only found for the item representing “quality of reflection” (item 6). The bivariate correlation coefficient for this item and the regular rating was 0.80, i.e. quality of reflection explained 64% of the variance in the regular end-of-year ratings ($p < 0.000$). The standard error of estimate - the measure of residual variance - was 0.41.

After entry of item 6, no further items accounted for a significant proportion of variance.

Table 3. Regression coefficient (R) and explained variance (R²) for the Inventory item on quality of reflection

Item	R	R ²	Standard error of estimate	Significance
Quality of reflection	0.803	64%	0.405	0.000

Conclusions

We examined which Inventory items affected portfolio assessment of reflective competence and which of these items carried the most weight. The results show that the ratings were primarily determined by quality of reflection. Quality of reflection was the only item to make a significant contribution (64%) to the explanation of the variance of the regular portfolio ratings. None of the other items of the Portfolio Quality Analysis Scoring Inventory made a substantial contribution.

The non-significant effects that we found for lay-out (item 1), spelling and grammar (item 2) and structure (item 3) as compared to quality of reflection (item 6) may be attributable to the mentoring and assessment training, which thus appears to positively affect the validity of the assessment procedure. Mentors and students meet twice a year and in those meetings mentors try to guide students' efforts into profitable directions. Mentors participate in an assessment training just before the end-of-year assessment. During this training session, mentors are given approximately 20 minutes to independently assess fragments of sample portfolios of previous years. These assessments are then compared for benchmarking and the mentors explain which criteria were decisive in determining their judgments. In this way it becomes clear which criteria are relevant and mentors learn to discriminate between relevant and irrelevant criteria.¹⁰

Apart from the mentor training, the portfolio template may be another factor that explains our findings.

A limitation of this study is the relatively small sample. Further research on a larger sample may reveal more significant regression coefficients. Another limitation is that the portfolio focused exclusively on reflective competence and did not include general clinical competencies.

Although we did not explicitly investigate this, the study appears to provide some support for the reliability of the assessment items as well. We found

rather high levels of agreement between 2 raters with regard to most items, not only the more objective ones, such as “The portfolio is complete”, but also more subjective items, such as “The student has looked critically at him/herself”. Other studies on portfolio assessment suggested that inter-rater reliability is probably enhanced when criteria are discussed by (2) independent raters, as was done in this study.^{16,17}

Because portfolio ratings were found to be associated with quality of reflection and not with aspects of presentation and writing style, we think we can conclude that the results support the validity of the global (holistic) assessment procedure for the assessment of reflective competence. Raters using the procedure appeared to be unaffected by irrelevant portfolio qualities in reaching their judgements.

References

1. McMullan M, Endacott R, Gray MA, Jasper M, Miller CML, Scholes J, Webb C. Portfolios and assessment of competence: a review of the literature. *J Adv Nurs* 2003;41(3):283-94.
2. Pitts J, Coles C, Thomas P. Educational portfolios in the assessment of general practice trainers: reliability of assessors. *Med Educ* 1999;33:515-20.
3. Cadbury-Amyot C, Kim J, Palm R, Mills E, Noble E, Overman P. Validity and reliability of portfolio assessment of competency in a baccalaureate dental hygiene program. *J Dent Educ* 2003;67(9):991-1002.
4. Tillema H. Design and validity of a portfolio instrument for professional training. *Studies Educational Eval* 1998;24(3):265-74.
5. Valencia S, Au K. Portfolios across educational contexts: issues of evaluation, teacher development, and system validity. *Educational Assessment* 1997;4(1):1-35.
6. Norman GR, Van der Vleuten CPM, De Graaff E. Pitfalls in the pursuit of objectivity: Issues of validity, efficiency and acceptability. *Med Educ* 1991;25:119-26.
7. Ryan JM, Kuhs TM. Assessment of preservice teachers and the use of portfolios. *Theory Pract* 1993;32:75-81.
8. Driessen EW, Van der Vleuten CPM, Schuwirth L, Van Tartwijk J, Vermunt JD. The use of qualitative research criteria for portfolio assessment as an alternative to reliability evaluation: a case study. *Med Educ* 2005;39:214-20.

9. Quinlan KM. Inside the peer review process: how academics review a colleague's teaching portfolio. *Teach Teach Educ* 2002;18:1035-49.
10. Heller J, Sheingold K, Myford C. Reasoning about evidence in portfolios: cognitive foundations for valid and reliable assessment. *Educational Assessment* 1998;5(1):5-40.
11. Driessen EW, Van Tartwijk J, Vermunt JD, Van der Vleuten CPM. Use of portfolios in early undergraduate medical training. *Med Teach* 2003;25(1):14-19.
12. Driessen EW, Van Tartwijk J, Overeem K, Vermunt JD, Van der Vleuten CPM. Conditions for successful use of portfolios for reflection. *Med Educ* 2005; 39:1230-35.
13. Cohen, J.A. Weighted Kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psych Bull* 1968;70: 213-20.
14. Graham P, Jackson R. The analysis of ordinal agreement data: beyond weighted kappa. *J Clin Epidemiol* 1993;46:1055-62.
15. Kianifard F. Evaluation of clinimetric scales: basic principles and methods. *Statistician* 1994;43:475-82.
16. Rees C, Sheard C. The reliability of assessment criteria for undergraduate medical students' communication skills portfolios: The Nottingham experience. *Med Educ* 2004;38:138-44.
17. Pitts J, Coles C, Thomas P & Smith, F. Enhancing reliability in portfolio assessment: discussions between assessors. *Med Teach* 2002; 24 (2):197-201.

6 Web-based or paper-based portfolios: Is there a difference?

**Erik W Driessen
Arno MM Muijtjens
Jan van Tartwijk
Cees PM van der Vleuten**

Published in: Medical Education 2007; 41: 1067-1073

Summary

Objective

To determine the differential effects of a paper-based portfolio versus a web-based portfolio in terms of portfolio quality, user-friendliness and student motivation.

Methods

An experimental design was used to compare Year 1 medical students' reflective portfolios. The portfolios differed in presentation medium only (i.e. web based versus paper based). Content analysis, a student questionnaire, and mentor interviews were used to evaluate portfolio quality, user-friendliness, and student motivation. A total of 92 portfolios were scored independently by 2 raters using a portfolio quality-rating instrument.

Results

Portfolio structure, quality of reflection and quality of evidence showed no significant effects of presentation medium. Multi-level analysis showed a significant effect for student motivation: web-based portfolios scored 0.39 more than paper-based portfolios ($P < 0.05$; effect size 0.76). The mentors reported no differences in portfolio quality, except for more visuals in web-based portfolios. Students spent significantly more time preparing the web-based portfolios than the paper-based ones (15.4 hrs. vs. 12.2 hrs.; $t = 2.1$, $P < 0.05$; effect size 0.46). The 2 student groups did not differ significantly in terms of satisfaction with the portfolio. The mentors perceived the web-based portfolios as more user-friendly.

Conclusions

The web-based portfolios were found to enhance students' motivation, were more user-friendly for mentors, and delivered the same content quality compared with paper-based portfolios. This suggests that web-based presentation may promote acceptance of portfolios by students and teachers alike.

Introduction

Portfolio use is on the increase in medical education. Portfolios are used not only as a source of information for assessments of functioning in authentic situations, but also to stimulate learners to reflect on their learning experiences.¹ Web-based portfolios (WBPs) are often preferred over paper-based portfolios (PBPs).^{2, 3} This preference is generally based on the following purported advantages of WBPs:

- hyperlinks in WBPs facilitate presentation and promote diversity of reflection and evidence (*portfolio quality*);³⁻⁵
- students are inclined to present their reflections in a more concise and well structured fashion in WBPs because the documents must be easy to read on screen; in many cases this improves both portfolio structure and readability (*portfolio quality*);⁶
- WBPs are easier to use and can be accessed by several persons simultaneously (*user-friendliness*);^{7,8} and
- students enjoy presenting themselves in a WBPs; WBPs enhance students' motivation and sense of ownership (*student motivation*).⁵

Thus, on the surface, WBPs appear to offer many advantages. It has been suggested that a possible drawback is that electronic presentation may diminish depth of reflection, because it may deflect students' attention from content to form.⁵ So far, however, there is little evidence to substantiate either the positive or the negative claims. What evidence is available is mostly confined to descriptions of individual portfolios or different types of portfolios,^{2,9} and measurements of students' and/or teachers' satisfaction with WBPs.^{7,10}

Usage of WBPs suffers from a problem, pinpointed by Cook, which is encountered with many computer applications in medical education: "...evidence supporting the use of these [web-based learning] tools is scant and often lags far behind technology, and there are some who fear that fascination with technology may outstrip actual learning gains."¹¹ As Cook suggested "...the time has come for hypothesis-driven comparative research in CBL and WBL" (i.e. in computer-based and web-based learning).¹¹ In view of the growing popularity of WBPs in medical education and elsewhere, it seems high time to gather some sound empirical evidence that can either support or refute claims that electronic portfolios offer added value compared to paper portfolios. The study we present in this paper was performed to gather such evidence. The key question of the study was: what are the

differential effects of a WBP versus a PBP regarding portfolio quality, user friendliness, and student motivation?

Methods

Design

We used an experimental design to compare WBPs and PBPs prepared by Year 1 medical students. The portfolios differed only in terms of the medium used (i.e. they were either web- or paper-based). Content, purpose, and procedure were identical. We compared the WBP with the PBP format on *portfolio quality, user-friendliness, and student motivation*.

Context

In the 6-year undergraduate medical curriculum of Maastricht University, the Netherlands, the learning environment in Year 1 is characterised by the use of authentic and real (patient) cases. The students in Year 1 use a portfolio to foster the development of their reflective skills. The design of this portfolio was based on experiences with portfolios described in the medical education literature and elsewhere.¹² Over the 4-year period since the introduction of this portfolio, we have refined the portfolio design guided by portfolio research conducted inside and outside our medical school.^{1,13}

The WBP and PBP that we compared in this study includes the same 3 components:

- written self-assessments of the student's personal development in 4 professional roles and the learning goals derived from these self-assessments, which are used to guide the student's ongoing development in these roles;
- evidence (artefacts, evaluations, etc.) underpinning the self-assessments, and
- feedback on the first 2 parts of the portfolio from the student's personal mentor.

The students have at least 2 one-on-one meetings with their personal mentor annually to discuss their portfolios. The mentor evaluates the student's reflective skills, as demonstrated in the portfolio, and suggests how these might be improved. This feedback is included in the portfolio.

On the basis of the portfolio, students' reflective skills are assessed (pass or fail) annually by the portfolio assessment committee. Although all mentors sit on the committee, they do not assess the portfolios of their own mentees.

The portfolio and the assessment procedure have been described in detail elsewhere.^{12,14}

Sample

Five of the total of 17 mentors in Year 1 of the undergraduate programme were randomly selected and asked to participate in this study. Participation was voluntary and each of the 5 mentors agreed to participate. Each mentor guides 2 groups of 9-10 students. One group of each mentor was randomly assigned to use the WBP ($n=45$) and one to using the PBP ($n=47$). The differences in sample size reflect the differences in mentor group size. All students gave informed consent for the use of their portfolios for research purposes.

Treatment

The WBP and the PBP that we compared were homogenous in configuration and instructional methods¹¹ (i.e. goals, structure, integration in the learning environment, instruction, coaching, and assessment). The only difference was the presentation medium:

- the PBP was presented in an A-4 ring binder with separate sections for the 4 professional roles, and
- the WBP was composed in the portfolio module of the medical school's electronic learning environment (Blackboard Content System), with a digital template for each of the professional roles.

Data Collection

A mixture of methods was used to evaluate portfolio quality, user-friendliness and student motivation (Table 1). We will briefly describe which research instrument provided information on which aspects of the research question.

Table 1. Study Design: Research Methods and Research Variables

	Portfolio quality	User-friendliness	Student motivation
Content analyses	WBP/PBP		WBP/PBP
Student questionnaire		WBP/PBP	WBP/PBP
Mentor interviews	WBP/PBP	WBP/PBP	

Content analysis

A slightly modified version of the Portfolio Quality Analysis Scoring Inventory¹⁵ was used for the content analysis of the portfolios. The Inventory was constructed in an earlier study and consists of 15 items derived from

interviews with experienced mentors and the literature.¹⁵ The analysis yielded information about *portfolio quality* and *student motivation*. For the present study, only the most relevant items of the instrument were used (table 2). The items were to be rated on a Likert scale of 1-5 (1 = definitely not applicable; 5 = definitely applicable).

Table 2. *Items of the Modified Portfolio Quality Analysis Scoring Inventory*

Quality of form and structure

- The portfolio is easy to use and information is easy to find

Quality of reflection

- The portfolio looks back to previous entries (i.e. what went wrong before, what went well this time and why, what did I say this time?)
- The analysis of strengths and weaknesses focuses on explanations for both (internal as well as external); the analysis goes beyond a mere listing of facts and situations
- The student has developed logical (following from the analysis of strengths and weaknesses) and clear learning goals in accordance with the SMART principle

Quality of evidence

- Whenever possible, statements are consistently supported by evidence
- Different types and sources of evidence are used

Additional effort put into the portfolio

- The student clearly made an effort to arrange the layout
 - The student has put more effort into portfolio content than was strictly required
-

Table 3. *Student questionnaire about the portfolio's user-friendliness and its effect on students' motivation to work on the portfolio*

1.	The portfolio is easy to use.
2.	I think it is a good thing that my mentor can access my portfolio via the Internet (WBP only)
3.	I enjoyed working on my portfolio.
4.	I am satisfied with my portfolio.
5.	Would you give a global estimate of the number of hours it took you to prepare your portfolio?

Students' and mentors' perceptions

We developed a short questionnaire to assess students' perceptions regarding *user-friendliness* and *student motivation* (Table 3). Items 1-4 were to be rated on a Likert scale of 1-5 (1=totally disagree, 5=totally agree).

Mentors' opinions about *user-friendliness* and *portfolio quality* were gathered in brief semi-structured interviews with the mentors. The topics addressed in each interview were:

- ease of use of the portfolio;
- differences in portfolio quality (form and content), and
- preference for portfolio medium.

Procedure

Each of the 92 portfolios was scored independently by 2 raters using the Portfolio Quality Analysis Rating Inventory. Three raters formed 3 different pairs and each pair scored approximately 61 portfolios. After rating 2 WBPs and 2 PBPs, the raters discussed the criteria in order to enhance inter-rater reliability.¹⁶ Thereafter, they rated the portfolios independently. Rater agreement was analysed using generalisability theory.¹⁷ With 2 raters, the dependability coefficient across items varied from 0.71 to 0.91. Of the 45 students in the WBP group and the 47 students in the PBP group, 42 (93%) and 42 (89%), respectively filled in the questionnaire after they had completed their Year 1 portfolios. Non-response reflected the absence of students in the week the questionnaire was administered. The first author (EWD) interviewed the 5 mentors and took notes.

Data analysis

Content analysis

To estimate whether there was an effect of the medium used (i.e. web versus paper), we investigated differences in *portfolio quality* and *student motivation* by analysing the rater pairs' mean ratings for question 1 (*Structure*), for question clusters 2-4 (*Quality of reflection*), 5 and 6 (*Quality of evidence*) and 7 and 8 (*Additional effort*). Because the portfolios were nested within mentors, the data had a multi-level structure, so the effect of the medium used was estimated by multi-level analysis. The analysis was based on a model defined on 2 levels, level 1 (Portfolio), and level 2 (Mentor), according to:

$$\text{level-1: } Score_{ij} = \beta_{0j} + \beta_{1j} * Medium + \varepsilon_{ij}$$

$$\text{level-2: } \beta_{0j} = \gamma_{00} + u_{0j}; \quad \beta_{1j} = \gamma_{10} + u_{1j}$$

where *Score* refers to the analysed rating; *Medium* is WBP or PBP (coded as 1 and 2); indices *i* and *j* refer to portfolio *i* and mentor *j*; β_{0j} is the level 1 intercept; β_{1j} is the difference in *Score* reflecting the different values of *Medium* (the effect of the medium used); u_{0j} , and u_{1j} represent (random) mentor effects for the intercept and the slope of the level 1 equation; γ_{00} , and γ_{10} are the fixed effects for the level 1 intercept and slope; and ε_{ij} is the level 1 residual. The computer program MLwiN version 1.10 was used for multi-level analysis.¹⁸

Students' and mentors' perceptions

T-tests were performed to estimate the differences in students' perceptions of *student motivation* and *user-friendliness*. Effect sizes were calculated and qualified according to Cohen.¹⁹ We compared the results of the mentor interviews using a cross case display matrix.²⁰

Results

Portfolio Quality

Content analysis

The results of the content analysis (5-point Likert scale) are shown in table 4. The multi-level analysis according to the model specified in Equation 1-2 showed a significant effect of portfolio medium for *Additional effort*: the WBP scored 0.39 higher than the PBP (joint Chi-square test, $P < 0.05$). The corresponding effect size is 0.76, which is considered large.¹⁹

Table 4. Mean scores (using a 5-point Likert scale), standard deviations (sd), standard errors (se) and number of respondents for the content analyses of the web-based and paper-based portfolios

	Web-based portfolio	Paper-based portfolio
Structure	4.3 (sd=0.75; se=0.11) N=45	4.0 (sd=0.77; se=0.11) N=47
Quality of reflection	4.1 (sd=0.58; se=0.09) N=45	4.1 (sd=0.68; se=0.10) N=47
Quality of evidence	4.2 (sd=0.83; se=0.12) N=45	4.1 (sd=0.73; se=0.11) N=47
Additional effort	3.1 (sd=0.47; se=0.07) N=45	2.7 (sd=0.54; se=0.08) N=47

For the clusters *Structure*, *Quality of reflection*, and *Quality of evidence*, no significant effects of the portfolio medium were found, although the WBP had considerably, albeit not significantly, higher scores for *Structure* compared with the PBP. No significant effects for mentor were found.

Mentors' perceptions

The mentors observed no differences between WBP and PBP with regard to content, except for the inclusion of more visuals in the WBPs.

User-friendliness

Mentors' perceptions

The 5 mentors were unanimous in their appraisal of the WBPs as easier to read. The reason they gave was that they did not have to leaf through the WBPs when they were looking for specific evidence. Hyperlinks made the WBPs easy to navigate. The mentors also indicated that looking for specific evidence took more time in the PBPs than in the WBPs. One mentor said:

“...in a paper portfolio I sometimes really have to search the appendices of the portfolio for the evidence the student refers to in the reflections. In the electronic portfolio I just click the hyperlink to go to the relevant evidence.” (Mentor 3)

Access from different locations was also mentioned as a positive point.

“I usually read the portfolios outside office hours. Therefore, it is nice that I can access the portfolios from my pc at home and don't have to carry the ringbinders between the faculty and home.” (Mentor 1)

Students' perceptions

The 2 student groups did not differ significantly in their evaluations of ease of use or enjoyment in working with the portfolios. The WBP group did not consider their format portfolio more user friendly than did the PBP group, but

they did consider it advantageous that their mentor could access the portfolio via the Internet (mean = 4.1, standard deviation = 1.0).

Student motivation

Students' perceptions

The WBP group spent significantly more time preparing their portfolios (WBP mean = 15.4 hours, SD = 7.6; PBP mean = 12.2 hours, SD = 5.7 hours; $t = 2.1$, $P = 0.05$). This difference is of moderate practical importance, with an effect size of 0.46, which is considered a medium-sized effect. The 2 groups did not differ significantly in terms of satisfaction with their respective formats.

Discussion

We investigated the effects of the medium used (i.e. web or paper), on *portfolio quality*, *user-friendliness* and *student motivation*. There was no significant difference between format for 3 out of 4 portfolio qualities. Any apprehension that WBPs might induce more superficial reflection was not confirmed by the results. However, the quality of the WBPs was not superior to that of the PBPs, so claims that students provide better and more diverse evidence in WBPs were equally unconfirmed. Both types of portfolio yielded a highly satisfactory quality of both reflection and supporting evidence. Structure differed, but not significantly so, although it was rated considerably more highly for WBPs than for PBPs.

The WBP scored significantly higher than the PBP on *Additional effort*. The students added more personal touches to the WBPs' content and form. Moreover, they spent significantly more time on the WBPs than on PBPs. This more substantial time investment in WBPs compared with PBPs cannot be explained by extra effort required for electronic presentation of the portfolio. The students were quite familiar with the electronic learning environment, and WBPs and PBPs take similar amounts of time to construct. Possibly, electronic presentation of the portfolio motivates the students, which incites them to spend more time on the portfolio. Several authors have stated that WBPs can have a motivating effect on students.^{5,21} Students are thought to enjoy presenting WBPs, which probably motivates them and enhances their sense of ownership.

The interviews showed that the mentors were unanimous in their appreciation of the greater ease of use of WBPs. They mentioned that information was

easy to locate without having to turn pages to find certain content and that the portfolios could be accessed from different locations. That is why the mentors prefer the WBP. Other authors have also reported the *user-friendliness* of electronic portfolios.^{7,8} Tutors appreciated the easy electronic access and reduction in the amount of paper.⁷ However, the same authors also reported about certain situations that made WBPs less user friendly than PBPs. For instance, limited computer access in the workplace cancels out the advantages of user-friendliness and could even have an opposite effect. The students in our study rated WBPs and PBPs equally in terms of user-friendliness.

What are the implications of the findings of this study? Ownership and motivation have been regarded as crucial factors for effective portfolio use.²² Students have generally reported that portfolios take up a great deal of time.²³ Moreover, students are generally not very keen to engage in self-reflection and do not readily do so of their own accord.^{24,25} In light of these reports, it seems important that the students in our study spent more time on the WBPs and showed a greater sense of ownership of the portfolio in this format with respect to both form and content. This suggests that web-based presentation may enhance acceptance of portfolios among students.

The increased user-friendliness for mentors can be viewed in the same light. Support by coaches has been shown to be a crucial factor in effective portfolio use.^{13,26} For teachers, reading portfolios can make a heavy claim on already limited time. The mentors in this study said that WBPs were easier and quicker to read than PBPs. We think this is a strong argument in favour of WBPs. With regard to content, there were no differences to support arguments in favour or against either type of portfolio. Apparently, the quality of portfolio content is not affected by the presentation medium.

Although we did not explicitly investigate this, the study appears to provide support for the reliability of the assessment of the quality of portfolios, since we found acceptable levels of agreement between the raters (0.71-0.91). Another interesting finding is that the mentors had no significant impact on portfolio quality. This may suggest that the mentor training and the peer assessment meetings that are regularly organised for mentors at Maastricht Medical School promote uniformity in the way mentors coach their students. A striking finding is that reflection was rated quite highly for both types of portfolio (4.1 on a 5-point scale). Very different experiences with reflective portfolios have been reported in the literature. Pearson and Heywood reported

that the majority of registrars and their trainers failed to use the portfolio for reflection, as was intended.²⁶ Most of the students in our study succeeded in producing a reflective portfolio of very satisfactory quality. This may be explained by the portfolio in our study meeting a number of conditions for effective portfolio usage, such as mentoring, a portfolio structure that stimulates reflection, and an adequate assessment procedure.¹³

A limitation of this study is that the students themselves were unable to compare between WBP and PBP, because each student used only 1 type of portfolio. Another limitation concerns the relatively small sample size of five mentors and 92 students for the survey part of the study (students' and mentors' perceptions). The results of the survey should therefore be interpreted with caution. Finally, portfolios are used for a variety of purposes in medical education. The portfolio examined in our study was expressly aimed at stimulating reflection in Year 1 medical students. A similar study of a portfolio aimed at assessing residents' clinical competence may yield very different outcomes.

In summary, the WBP enhances student motivation, improves ease of use for mentors and delivers quality of portfolio content that is equal to that of the PBP. These outcomes of our comparison between a web-based and a paper-based portfolio appears to warrant an overall conclusion in favour of the web-based format.

References

1. Carraccio C, Englander R. Evaluating competence using a portfolio: a literature review and web-based application to the ACGME competencies. *Teach Learn Med* 2004;16(4):381-7.
2. Duque G. Web-based evaluation of medical clerkships: a new approach to immediacy and efficacy of feedback and assessment. *Med Teach* 2003;25(5):510-4.
3. Campbell C, Parboosingh J, Gondocz T, Babitskaya G, Pham B. A study of the factors that influence physicians' commitments to change their practices using learning diaries. *Acad Med* 1999;74(10 suppl):S34-S36.
4. Kimeldorf M. *Portfolio Power: the New Way to Showcase all your Job Skills and Experience*. Princeton NJ: Peterson's Publishing Group, 1997.
5. Woodward H, Nanlohy P. Digital portfolios: fact or fashion? *Assess Eval Higher Educ* 2004;29(2):227-38.

6. Price J, Price L. *Hot Text, Web Writing that Works*. Berkley, CA: New Riders 2002.
7. Lawson M, Nestel D, Jolly B. An e-portfolio in health professional education. *Med Educ* 2004;38:569-70.
8. Fung Kee Fung M, Walker M, Fung Kee Fung K, Temple L, Lajoie F, Bellemare G, et al. An Internet-based learning portfolio in resident education: The KOALA-super(TM) multicentre programme. *Med Educ* 2000;34(6):474-9.
9. Supiano M, Fantone J, Grum C. A web-based geriatrics portfolio to document medical students' learning outcomes. *Acad Med* 2002;77(9):937-8.
10. Dornan T, Lee C, Stopford A. Skills base: a web-based electronic learning portfolio for clinical skills. *Acad Med* 2001;76(5):542-3.
11. Cook DA. The research we still are not doing: an agenda for the study of computer-based learning. *Acad Med* 2005;80(6):541-8.
12. Driessen EW, Van Tartwijk J, Vermunt JD, Van der Vleuten CPM. Use of portfolios in early medical training. *Med Teach* 2003;25(1):14-19.
13. Driessen EW, Van Tartwijk J, Overeem K, Vermunt JD, Van der Vleuten CPM. Conditions for successful use of portfolios for reflection. *Med Educ* 2005;39:1230-5.
14. Driessen EW, Van der Vleuten CPM, Schuwirth LW, Van Tartwijk J, Vermunt JD. The use of qualitative research criteria for portfolio assessment as an alternative for reliability evaluation: a case study. *Med Educ* 2005;39:214-20.
15. Driessen EW, Overeem K, Van Tartwijk J, Van der Vleuten CPM. Validity of portfolio assessment: which qualities determine ratings? *Med Educ* 2006; 40: 862-866.
16. Rees C, Sheard C. The reliability of assessment criteria for undergraduate medical students' communication skills portfolios: The Nottingham experience. *Med Educ* 2004;38:138-44.
17. Brennan RL. *Generalizability theory*. New York: Springer-Verlag 2001.
18. Rasbash J, Browne W, Goldstein H, Yang M, Plewis I, Healy M, et al. *A User's Guide to MLwiN*. London: Centre for Multilevel Modelling, Institute of Education, University of London 2000.
19. Cohen J. *Statistical Power Analysis for Behavioral Sciences*. Hillsdale, NJ: Erlbaum 1987.
20. Miles MB, Huberman M. *Qualitative Data Analyses: an Expanded Source Book*, 2nd ed. Thousand Oaks CA: Sage Publications, 1994.

21. Gibson D, Barrett H. Directions in electronic portfolio development. *Contemp Issues Technol Teach Educ* 2003;2(4):559-76.
22. Snadden D, Thomas ML. Portfolio learning in general practice vocational training - does it work? *Med Educ* 1998;32:401-6.
23. Davis M, Friedman Ben-David M, Harden R, Howie P, Ker C, McGhee C, Pippard M, Snadden D. Portfolio assessment in medical students' final examinations. *Med Teach* 2001;23:357-66.
24. Dornan T, Maredia N, Hosie L, Lee C, Stopford A. A web-based presentation of an undergraduate clinical skills curriculum. *Med Educ* 2003;37:500-8.
25. Grant A, Kinnersley P, Metcalf E, Pill R, Houston H. Students' views of reflective learning techniques: an efficacy study at a UK medical school. *Med Educ* 2006;40:379-88.
26. Pearson DJ, Heywood P. Portfolio use in general practice vocational training: a survey of GP registrars. *Med Educ* 2004;38:87-95.

7 Portfolios in medical education: why do they meet with mixed success? A systematic review

**Erik W Driessen,
Jan van Tartwijk,
Cees PM van der Vleuten
Val J Wass**

Published in: Medical Education 2007; 41: 1224-1233

Summary

Context

The move towards competence-based medical education has created a need for instruments to support and assess competence development. Portfolios seem suitable but mixed reports of their success are emerging.

Methods

To examine the effectiveness of portfolios, we searched PubMed and EMBASE using the keyword “portfol*”, Psychinfo and ERIC using the keywords “portfol*” and “medical education” and references of retrieved papers for empirical studies on portfolios in all phases of medical education. Thirty of 1939 retrieved papers met the inclusion criteria and were analysed. Data were collated against the research question, number of subjects, design, setting, findings and limitations, purpose and content, mentoring and assessment. We analysed impact using a modified version of Kirkpatrick’s hierarchy.

Results

Because differences across studies precluded statistical meta-analysis. the data were analysed by context, goals, and procedure. Positive effects were strongest in undergraduate education. Important factors for success were: clearly communicated goals and procedures; integration with curriculum and assessment; flexible structure; support through mentoring, and measures to heighten feasibility and reduce required time. Moderately good inter-rater reliability was reported and global criteria and discussions among raters were beneficial. Formative and summative assessment could be combined. Without assessment, portfolios were vulnerable to competition from other summative assessment instruments.

Conclusions

For portfolios to be effective in supporting and assessing competence development, robust integration into the curriculum and tutor support are essential. Further studies should focus on the effectiveness and user-friendliness of portfolios, the merits of holistic assessment procedures, and the competences of an effective portfolio mentor.

Introduction

Over the last 2 decades, a significant change has occurred in medical education. The focus of curricula has shifted from the acquisition of knowledge to the achievement of competence.^{1,2} Competence has been defined as “the habitual and judicious use of communication, knowledge, technical skills, clinical reasoning, emotions, values, and reflection in daily practice for the benefit of the individuals and communities being served.”³ The challenge has been to find instruments that formatively support the development of competence in an integrated, coherent and longitudinal fashion and summatively assess whether competence is being achieved.^{4,5} Portfolios are acclaimed as such an instrument.⁶ The past 10-15 years have seen the introduction of portfolios in all stages of the medical education continuum: in undergraduate medical education;^{5,7} postgraduate specialist training,⁸⁻¹⁰ and continuing medical education (CME) of practising doctors.^{6,11-14}

Portfolios that are used in education contain evidence of how trainees fulfil tasks and their competence is progressing. Portfolios may be digital or paper-based and content may be prescribed or left to the students’ discretion. Despite variations in content and format, portfolios basically report on work done, feedback received, progress made, and plans for improving competence. Additionally, portfolios may stimulate reflection, because collecting evidence for inclusion in a portfolio requires looking back and analysing what one has accomplished.

Reflection can be defined as the mental process of trying to structure or restructure an experience, a problem, or existing knowledge.¹⁵ This can help learners to understand their development¹⁶ and plan their learning.¹⁵ Reflecting on task performance and development of competence implies self-assessment or self-rating;¹⁷ learners have to compare their own performance to (external) standards. Reflection and self-assessment are essential skills for lifelong learning, but the literature on self-assessment is quite clear in showing that students and physicians have a limited ability to self-assess their competence and learning needs.¹⁷ Hence, it has been suggested that self-assessment should be supported by other (external) sources of information.^{17,18} Portfolios may have the potential to improve self-assessment, by combining external assessment, mentoring and self-assessment.¹⁹

Since their introduction in medical education in the early 1990s, portfolios have been the subject of educational research. The evidence to date suggests that their introduction has met with mixed success.²⁰⁻²² There is little explanation for these differences, which could relate to many factors. The aim of this study was to conduct a systematic review of the literature on portfolios to seek evidence and clarify why in some contexts portfolios appear to be largely ineffective, whereas in others they are successful.

Methods

Data sources

Pubmed (1966-May 2007) and EMBASE (1989-May 2007) were searched using the keyword “portfol*”. The databases Psychinfo (1970-May 2007) and ERIC (1966-May 2007) were searched using the keywords “portfol*” and “medical education”.

The searches were limited to publications in English and Dutch, because it was not feasible to translate non-English or non-Dutch articles. To identify studies not picked up in the initial search, we contacted experts in the field and checked the references of the papers retrieved by the initial search.

Selection of studies

We used broadly defined inclusion criteria to ensure all aspects of the research question were addressed. We included studies that:

1. focused on portfolio use for educational purposes in medical training;
2. were performed within the context of undergraduate, postgraduate, or CME, and
3. reported empirical data.

We excluded studies concerning: portfolios for other health professions (nursing, dentistry, dietetics, veterinary medicine), administrators, managers, teachers, and trainers in hospitals, management, finance, education, teaching, specialist trainers and academic portfolios; portfolio-related instruments, such as logbooks, personal digital assistants, and personal development plans and descriptive articles without evaluative data.

Data abstraction

The literature search was performed by 2 of the authors (ED, JvT) and an information specialist. Three of the authors (ED, JvT, CvdV) determined the inclusion criteria. Two of the authors (ED, JvT) supported by a third author

(CvdV), reviewed the titles and abstracts of retrieved publications and selected relevant articles for possible inclusion. Data abstraction methods were developed by 3 of the authors (ED, JvT, CvdV) and were applied by 2 of the authors (ED, JvT). Disagreements about search criteria, data abstraction, and classification of study results were resolved by consensus. The reviewers were not blinded to any portion of articles. The authors of 1 of the studies were contacted and asked for clarification of some points, which they did.

The articles fulfilling the selection criteria referred to a wide range of studies where portfolios were used for different purposes in a variety of contexts within medical education, and methods and quality varied. Most of the selected studies used a variety of measurement methods and surveyed a range of portfolios which differed in purpose, content and format. With the exception of inter-rater reliability, statistical pooling of the results proved impossible. We made narrative descriptions of the findings and quality of the studies according to the criteria suggested by the Best Evidence Medical Education Collaboration (BEMEC).²³ Assessment of the quality of the studies was based on the study design, questionnaire validation, sampling frame and size, response rate, and outcome measures.²⁴ Data were synthesised and reported where possible in relation to the influence of context and portfolio goals to address the aims of the review.

We used a modified version of the BEMEC coding sheet for data abstraction.²³ The form included details of the research question, number of subjects, study design, setting, findings, and study limitations. Further details of the intervention (i.e. the goal, contents and structure, and the mentoring and assessment of the portfolio²⁵) were also included. The impact of the intervention was rated using a modified version of Kirkpatrick's hierarchy to analyse outcomes such as learner satisfaction, learning outcomes, performance improvement, and patient or health outcomes.²⁶

The inter-rater reliability across all studies was estimated by averaging domain-referenced reliability coefficients or kappas. The Spearman-Brown prophecy formula was used to estimate the projection of inter-rater reliability for the use of multiple raters.²⁷

Results

Search results

The search revealed 1939 publications. After reading titles and abstracts, we excluded 1853 articles which failed to meet the inclusion criteria. More detailed review of the remaining 86 publications yielded 30 articles that met the inclusion criteria (Fig. 1). Of these articles, 9 related to the use of portfolios in preclinical undergraduate medical education, 7 addressed undergraduate clinical clerkships, 9 concerned postgraduate medical training, and 5 dealt with CME (see tables 1-4 in the appendix of this chapter).

Many of these studies had methodological limitations. With the exception of 2 studies,^{21,28} all had a single-group design. The majority was conducted in a single institution. In 5 studies, participants were self-selected volunteers.^{20,28-31} Many studies lacked a detailed description of the portfolio, how it was introduced to its users, the sampling frame, the study method, data analysis or outcomes, which limited our ability to fully appraise the quality of the study or generalise the findings.

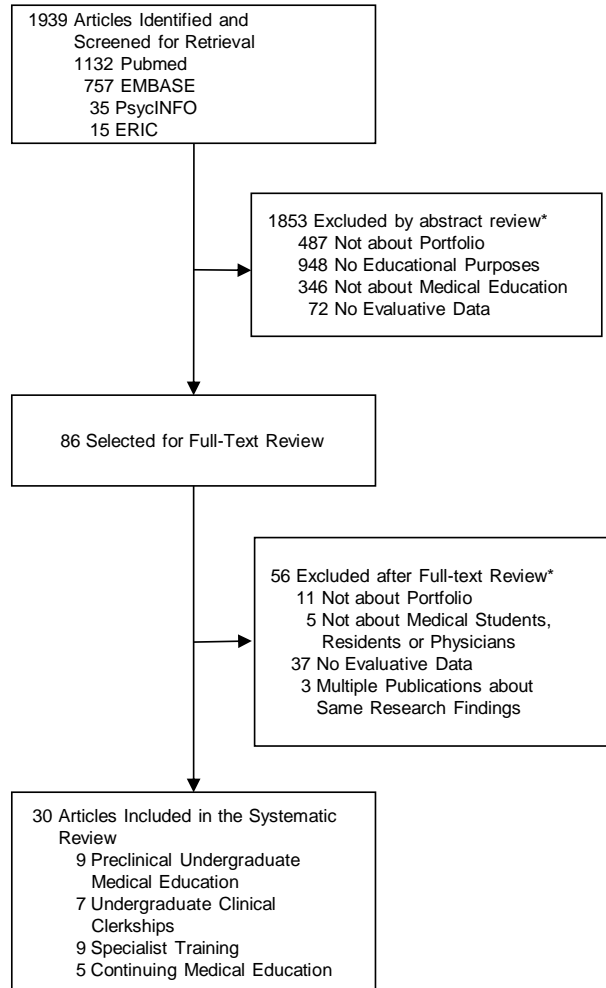
A total of 19 studies evaluated outcomes at Kirkpatrick level 1 (i.e. surveying the satisfaction level of the users). Only 2 studies reported outcomes in terms of performance improvement (level 3).^{22,30} None of the studies measured patient outcomes.

We report the results in relation to the 2 broad portfolio goals: learning and assessment.

Goal 1. Learning

Two studies performed in the preclinical phase of medical school reported that portfolios contributed to Year 1 students' reflective learning^{32,33} and 1 study reported that portfolios contributed to students' personal and professional development.²² Two studies reported successful use of portfolios to organise, monitor and evaluate a pre-clinical oncology programme and help students understand the impact of malignant disease on patients.^{21,34} Studies where portfolios were used in CME yielded mixed results. Some reported that portfolios could stimulate reflective learning^{28,30} and support planning and monitoring of CME.³⁰

Figure 1. Selection of articles for review



*Some articles excluded for multiple reasons

Many studies across a range of contexts reported problems related to poor preparation and introduction of portfolios by the institution. Examples of this claimed either that the purpose of the portfolio was not clearly defined³⁵ or that learners and teachers were poorly or insufficiently informed about the portfolio and what it entailed.³⁵⁻³⁸ In 1 study this resulted in poor commitment from both residents and their trainers and limited use of the portfolio.³⁵ However, in 5 studies where portfolio design centered on informing, training and gaining commitment from both trainers and trainees, portfolios were found to be suitable for graduate training.³⁹⁻⁴³ One study demonstrated that

hands-on introduction with a proper briefing of students by the staff on the portfolio's purpose and procedures, had a positive effect on portfolio scores and students' satisfaction with the portfolio.⁴⁴

The use of portfolio in undergraduate education was more successful when portfolios were not used in isolation but were part of other educational activities.^{21,22,32,34} These educational activities included pairing students with oncology patients,^{21,34} organising tutorial groups³⁵ or mentoring,³² or linkage to an interview.²²

Results of the use of portfolios in CME often suggested that portfolios were not used by doctors to their full potential. This was related mainly to time constraints imposed by high daily workloads^{20,31} and the perception that the portfolio was timeconsuming.^{20,28,45} On occasion studies referred to extraneous issues, such as difficulties with information technology, for instance problems with downloading necessary software²⁰ or lack of IT skills.³¹ Lack of time was also an issue for postgraduate training.^{38,41,46} Trainees and their supervisors were concerned that the portfolio might be too time intensive and for this reason avoided using it.^{38,41,46} Mathers *et al.* made a plea for portfolios to be “smarter” (less paperwork) to aid feasibility.²⁸ A study investigating the use of such an efficient portfolio supports this supposition, as undergraduate clerks did not find the portfolio labour-intensive.⁴⁷

The format of the portfolio also influenced the contribution it made to learning. An effective portfolio had a clear but flexible structure, allowing the learners the opportunity to describe their own unique development.^{32,35,48} Clear instructions were important. Most users wanted to know what kind of information they were expected to provide.^{29,44,49} In clinical contexts where the content of a portfolio was often highly prescribed, portfolios were experienced as bureaucratic instruments.^{35-37,45,46} Portfolios were more highly appreciated when learners were given a certain amount of freedom to determine their content.^{32,50}

Many studies reported the lack of adequate support from mentors.^{20,31,35} Other studies confirmed that mentoring by teachers, trainers or educational supervisors made an important contribution to the success of the portfolio.^{21,28,32,37,39-41,48} Mentors included teachers, trainers, supervisors, or peers.³¹ General Practice trainees made more use of their portfolios when they had a supportive trainer.^{35,38,41} Because of the significant impact of

mentoring, it was difficult sometimes to discriminate between the effects of the mentor and the practicalities of completing the portfolio itself.³⁵ Obviously, mentoring requires teacher and supervisor time.³² However mentoring aimed at stimulating the development of reflective ability³² and deep learning strategies focused on comprehension and understanding^{28,43} merited the effort.

Goal 2: Assessment

A study investigating the validity of portfolio assessment⁵¹ demonstrated it was indeed a valid test of reflective ability. Quality of reflection was the strongest predictor of the final assessment grade. Other criteria, such as layout and writing style, had a negligible effect. Six studies estimated the inter-rater reliability of portfolios.^{45,46,48,51-53} The average reliability across these 6 studies was 0.63, representing the estimated reliability if one assessor would be randomly replaced by another one. However, with 2, 3 or 4 raters, the reliability would increase to 0.77, 0.84, and 0.87, respectively. A value of 0.80 is usually required for high-stakes tests.²⁷ The studies suggested that a number of measures had a positive impact on inter-rater agreement: use of a small group of (trained) assessors;^{45,46,51-53} discussion among the raters before the actual assessment^{46,51-53} and after assessing part of the portfolio,^{46,53} and the use of global criteria with rubrics.^{45,46,53}

In general, there was more support for the formal assessment of portfolios from teachers and examiners than from the students themselves.^{36,44,45} For example, in a study in which examiners were positive about the use of portfolios for assessment, final year undergraduate students reported that a comprehensive portfolio with prescribed content involved far too much paperwork and, if they were to be formally assessed, they needed more advance information about how to construct the portfolio.³⁶

The use of portfolio for assessment and learning is often seen as conflicting: students may be less open in their reflections when their portfolios are assessed. However, 2 studies examining the combination of formative mentoring and summative assessment in 1 portfolio reported that this was not an issue. In 1 study mentors reported that portfolio assessment had no effect on students' openness³² and for general practitioners and their CME tutors the combination of support and assessment did not appear to be problematic.²⁸ Two studies showed that if portfolios were not formally assessed, other summative assessment instruments became the priority and the use of portfolios tailed off.^{35,43}

Discussion

To our knowledge, this is the first systematic review of the literature on portfolio use in medical education. We found many descriptive articles, opinion papers, and commentaries on portfolio use. Only 30 of the retrieved articles reported empirical data. The available evidence demonstrated that portfolios can support both the learning and the assessment of more general, yet essential, competences in pre-clinical undergraduate education, such as reflective ability, personal and professional development, communication skills, and empathy towards terminally ill patients and their families. This finding is consistent with the evidence from a recent literature review of portfolio use in nursing.⁵⁴ Portfolios also have potential as a tool to organise workplace learning during clerkships and postgraduate specialist training. Here, a more mixed picture emerged of contrasting poor and successful examples of portfolio introduction at all stages of training. Our review is in agreement with earlier literature, showing that several key issues are decisive in the successful use of a portfolio⁵⁵ and is consistent with studies in other disciplines. For example, in 1996 a study in teacher education showed that, for a portfolio to stimulate reflection, certain conditions had to be met, i.e. a good introduction of the portfolio and its intended use, student ownership, a clear structure, and appropriate use of the portfolio in discussions with mentors or trainers.⁵⁶ McMullan concluded in a recent study in nursing that portfolios can be very effective as an assessment and learning tool, but only if both students and mentors receive clear guidelines and support for their use.⁵⁷ McMullan noted, that without support and clear guidelines, students and mentors became increasingly stressed and demoralised with the use of portfolios in practice.⁵⁷ The studies highlight several success factors for portfolio use.

Success factors

The goals of working with a portfolio need to be clear but can be successfully combined.^{28,32} Portfolios can be used concurrently both to formatively promote learning and for summative assessment. This is in contrast to previous debate in the literature, where the use of portfolio for assessment and learning was seen as conflicting: students may be less open in their reflections when their portfolios are assessed.⁵⁸⁻⁶⁰ From this systematic review of the literature, combining the 2 goals of learning and assessment does not appear to cause problems. On the contrary, summative assessment was important to ensure that portfolio learning maintained its status alongside other assessed subjects.^{35,43}

It is advisable to regard a portfolio not as a separate, independent instrument, but to integrate it with other educational activities.^{21,22}

Effectiveness of learning is enhanced by providing a mentor to support the portfolio. Mentorship requires a substantial time investment but appears crucial for successful use of portfolio.^{21,28,32,35,37,39,41,43,48} The effectiveness of assessment can be enhanced by combining the portfolio with an interview.^{22,36,47}

A major challenge for the integration of a portfolio into medical education is to maintain its status in the eyes of assessment-driven students. This review suggests that it must be part of the institutional assessment procedures.^{32,35,43} We found surprisingly high levels of inter-rater reliabilities in the studies.^{45,46,48,51-53} This contrasts with findings in other domains, such as that of teaching portfolios of General Practice trainers.^{61,62} The results of our review suggest that assessment panels may be limited to 2 or 3 assessors depending on the stakes of the assessment. Part of the success in achieving high reliabilities appears to be attributable to the use of a small group of trained assessors,^{45,46,51-53} specific assessor training exercises,^{46,51-53} including benchmarking, assessor discussion (before and intermediate) and use of holistic scoring rubrics (global performance descriptors).^{45,46,53} In her review of portfolio assessment in nursing, McCready⁵⁴ also calls for experienced assessors, explicit guidelines for portfolio construction and a holistic assessment procedure. The good news seems to be that with these appropriate measures adequate assessment of portfolios is possible, without the need to prescribe the content and structure of the portfolio in detail.⁶³

Another issue that impacts on portfolio success is a flexible learner-centered format. A rigid structure in which every detail of portfolio content is prescribed elicits negative reactions from portfolio users and is regarded as counterproductive.^{28,32,35,36} Too much structure appears from this review and other literature^{56,57} to carry a greater risk than too little structure. This does not deny the fact that learners do need clear directions and guidance to support the development and assessment of broad competencies.^{29,57} However, direction should be achieved through clear guidelines and well-defined portfolio goals rather than minute directives for every detail of the portfolio.⁵⁷ Striking the right balance is crucial here.

Time or rather lack of it, is another key issue. Many learners who are asked to create a portfolio, and their supervisors/mentors, are concerned that building and judging portfolios will be exceedingly time consuming or downright

impossible.^{20,28,31,38,41,45,46,64} The finding that time constraints appear to be less of a problem for pre-clinical students may reflect that students had relatively more time at their disposal. For learners in clinical settings, it is clearly difficult to find time amidst the pressures of clinical practice. Many of the portfolios described in the studies we reviewed were not user friendly and involved huge amounts of paperwork, forcing portfolio users to comply with strict and detailed guidelines.^{35,36,45,46} Too much obligatory content in particular makes portfolios bureaucratic, both failing to serve any educational purpose and forcing learners to search for content outside their direct and lived experiences.^{32,57}

Study limitations

Several limitations in this review should be considered. Firstly, the label “portfolio” refers to a broad range of instruments. The purpose, context, structure and content of portfolios that were used, differed considerably. Because of these differences and the variety in study methods and study quality, it was not possible to use a statistical meta-analytic approach. We attempted to overcome this limitation by synthesising the data as much as possible per context and per goal. Secondly, the literature in medical education often lacks the use of extensive medical subject headings, which could have contributed to not retrieving some studies. In addition, different labels were sometimes used in the text of articles. Because we excluded studies of instruments like logbooks, appraisals, personal digital assistants, and personal development plans, we may have missed studies in which they were similar to portfolios if the authors did not use the term “portfolio”. Although we manually searched reference lists to overcome these subject heading and label limitations, we may have missed some studies. Thirdly, many studies lacked a full description of the actual portfolio, the portfolio introduction, the study method, data analysis, and outcomes. This limited our ability to describe the studies more fully or to generalise more. Finally, in some studies it was not possible to differentiate if the observed outcomes were the result of working with a portfolio or of mentoring³⁵ or other educational activities.^{21,22,28} We believe, however, that future studies should not try to solve this limitation, as this review showed the crucial importance of integrating portfolios and mentoring in the curriculum.

Implications for research and practice

The results of this review show that many questions regarding portfolio use are still unanswered and this has important implications for both research and practice. We found many studies where the description of the portfolio structure and its implementation were inadequate. In view of the wide

variation in portfolio formats, researchers and peer reviewers should insist that details of portfolio structure (purpose, content, mentoring and assessment)²⁵ are given, along with the context in which it was implemented, to ensure others can critically appraise their papers adequately. Although the literature indicates that portfolios are not always successful, many studies did not examine how they were implemented and why they failed. We found no studies investigating the influence of the context in which a portfolio is introduced. To claim success for an educational intervention, such as portfolios, researchers need to look carefully at the intervention in practical settings.⁶⁵

The implementation of a portfolio requires greater rigour than we encountered in many papers. This lack of scientific rigour may account for our disappointing finding that there was no trend toward improvement of portfolio delivery over the time span represented by these studies.

Future portfolio research could focus on the user-friendliness or feasibility of portfolios and address time constraints, (for example portfolios that are supported by curriculum arrangements, such as protected time for learning),⁶⁶ the merits of holistic assessment procedures, and the competences of effective portfolio mentors.

References

- 1 Rabinowitz HK, Babbott D, Bastacky S, et al. Innovative approaches to educating medical students for practice in a changing health care environment: The National UME-21 project. *Acad Med* 2001;76:587-595.
- 2 National Health Service. *Modernising Medical Careers Foundation Programmes*. <http://www.mmc.nhs.uk/pages/foundation>. [Accessed August 29, 2006.]
- 3 Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* January 9 2002;287(2):226-235.
- 4 Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* September 1990;65(9 Suppl):S63-67.
- 5 Epstein RM. Assessment in medical education. *N Engl J Med* 2007;356:387-396.
- 6 Royal College of General Practitioners. *Portfolio-based learning in general practice: Report of a working group on higher professional education: December 1993*. Occasional paper 63.

- 7 General Medical Council. *Tomorrow's doctors: Recommendations on undergraduate medical education*. <http://www.gmc-uk.org/education/undergraduate/tomdoc.pdf>. [Accessed August 29, 2006.]
- 8 Accreditation Council for General Medical Education, American Board of Medical Specialties. *Toolbox of Assessment Methods*, version 1.1. <http://www.acgme.org/Outcome/assess/Toolbox.pdf>. [Accessed August 29, 2006.]
- 9 Hayden SR, Dufel S, Shih R. Definitions and competencies for practice-based learning and improvement. *Acad Emerg Med* Nov 2002;9(11):1242-1248.
- 10 Lynch DC, Swing SR, Horowitz SD, Holt K, Messer JV. Assessing practice-based learning and improvement. *Teach Learn Med* 2004;16(1):85-92.
- 11 Alliance for Academic Internal Medicine. *Portfolio for Internal Medicine Residency Programs*. <http://www.im.org/AAIM/Tools/Docs/Forms/competencies.pdf>. [Accessed August 29, 2006.]
- 12 Bennett NL, Davis DA, Easterling WE Jr., et al. Continuing medical education: A new vision of the professional development of physicians. *Acad Med* 2000;75:1167-1172.
- 13 Lewis M, Elwyn G, Wood F. Appraisal of family doctors: an evaluation study. *Br J Gen Pract* 2003;53:454-460.
- 14 Zeiger RF. Toward continuous medical education. *J Gen Int Med* 2005;20(1):91-94.
- 15 Korthagen FAJ, Kessels J, Koster B, Lagerwerf B, Wubbels T. *Linking theory and practice: The pedagogy of realistic teacher education*. Mahwah, NY: Lawrence Erlbaum Associates; 2001.
- 16 Klenowski V. *Developing portfolios for learning and assessment: Processes and principles*. London: Routledge/Falmer; 2002.
- 17 Davis DA, Mazmanian PE, Fordis M, Van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* Sep 6 2006;296(9):1094-1102.
- 18 Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med* Oct 2005;80(10 Suppl):S46-54.
- 19 Van Tartwijk J, Driessen E, Stokking K, Van der Vleuten C. Factors influencing the successful introduction of portfolios. *Qual Higher Educ* 2007;13(1):69-79.

- 20 Dornan T, Carroll C, Parboosingh J. An electronic learning portfolio for reflective continuing professional development. *Med Educ* 2002;36(8):767-769.
- 21 Finlay IG, Maughan TS, Webster DJ. A randomized controlled study of portfolio learning in undergraduate cancer education. *Med Educ* 1998;32(2):172-176.
- 22 Gordon J. Assessing students' personal and professional development using portfolios and interviews. *Med Educ* 2003;37(4):335-340.
- 23 Best Evidence Medical Education Collaboration. *Information for Prospective Topic Review Groups*. <http://www.bemecollaboration.org/TRGInfo.htm>. [Accessed June 19, 2006.]
- 24 Reed D, Price EG, Windish DM, et al. Challenges in systematic reviews of educational intervention studies. *Ann Intern Med* 2005 Jun 21;142:1080-1089.
- 25 Zeichner K, Wray S. The teaching portfolio in US teacher education programs: What we know and what we need to know. *Teach Teach Educ*. 2001;17:613-621.
- 26 Curran VR, Fleet LA. A review of evaluation outcomes of web-based continuing medical education. *Med Educ* 2005;39:561-567.
- 27 Magnusson D. *Test Theory*. Stockholm: Addison-Wesley; 1967.
- 28 Mathers NJ, Challis MC, Howe AC, Field NJ. Portfolios in continuing medical education - effective and efficient? *Med Educ* 1999;33(7):521-530.
- 29 Vink S, Cools HJM, Elsen GMF, Vermunt JD. Reflectie van eerstejaars studenten geneeskunde: wat willen we zien en wat doen de studenten? (in Dutch with summary in English) *Dutch J Med Educ* 2005;24(1):4-13.
- 30 Campbell CM, Parboosingh JT, Gondocz ST, et al. Study of physicians' use of a software program to create a portfolio of their self directed learning. *Acad Med* 1996;71(10):S49-50.
- 31 Dagley V, Berrington B. Learning from an evaluation of an electronic portfolio to support general practitioners' personal development planning, appraisal and revalidation. *Educ Prim Care* 2005;16:567-574.
- 32 Driessen EW, Van Tartwijk J, Overeem K, Vermunt JD, Van der Vleuten CPM. Conditions for succesful reflective use of portfolios in ndergraduate medical education. *Med Educ* 2005;39:1230-1235.
- 33 Driessen EW, Van Tartwijk J, Vermunt JD, Van der Vleuten CPM. Use of portfolios in early undergraduate medical training. *Med Teach* 2003;25(1):14-19.

- 34 Maughan TS, Finlay IG, Webster DJ. Portfolio learning with cancer patients: An integrated module in undergraduate medical education. *Clin Oncol* 2001;13(1):44-49.
- 35 Pearson DJ, Heywood P. Portfolio use in general practice vocational training: A survey of GP registrars. *Med Educ* 2004;38(1):87-95.
- 36 Davis MH, Friedman BDM, Harden RM, et al. Portfolio assessment in medical students' final examinations. *Med Teach* 2001;23(4):357-366.
- 37 Elango S, Jutti RC, Lee LK. Portfolio as a learning tool: students' perspective. *Ann Acad Med Singapore* 2005;34(8):511-514.
- 38 Kjaer NK, Maagaard R, Wied S. Using an online portfolio in postgraduate training. *Med Teach* 2006;28(8):708-712.
- 39 Ryland I, Brown J, O'Brien N, et al. The portfolio: how was it for you? Views of F2 doctors from the Mersey Deanery Foundation Pilot. *Clin Med* 2006; 6(4):378-380.
- 40 Webb TP, Aprahamian C, Weigelt JA, Brasel KJ. The Surgical Learning and Instructional Portfolio (SLIP) as a self-assessment educational tool demonstrating practice-based learning. *Curr Surg* 2006;63(6):444-447.
- 41 Snadden D. Portfolio-based learning and general practice vocational training. *Med Educ* 1996;30:148-152.
- 42 Fung MKF, Walker M, Fung KKF, et al. An Internet-based learning portfolio in resident education: The KOALA-super(™) multicentre programme. *Med Educ* 2000;34(6):474-479.
- 43 Snadden D, Thomas ML. Portfolio learning in general practice vocational training - does it work? *Med Educ* 1998;32(4):401-406.
- 44 Duque G, Finkelstein A, Roberts A, Tabatabai D, Gold SL, Winer LR. Learning while evaluating: the use of an electronic evaluation portfolio in a geriatric medicine clerkship. *BMC Med Educ* 2006;6:4.
- 45 Tate P, Foulkes J, Neighbour R, Champion P, Field S. Assessing physicians' interpersonal skills via videotaped encounters: A new approach for the Royal College of General Practitioners Membership examination. *J Health Commun* 1999;4(2):143-152.
- 46 O'Sullivan PS, Reckase MD, McClain T, Savidge MA, Clardy JA. Demonstration of portfolios to assess competency of residents. *Adv Health Sci Educ* 2004;9(4):309-323.
- 47 Burch VC, Seggie J. Portfolio assessment using a structured interview. *Med Educ* 2005;39(11):1169.
- 48 Grant AJ, Vermunt JD, Kinnersley P, Houston, H. Exploring students' perceptions on the use of significant event analysis, as part of a

- portfolio assessment process in general practice, as a tool for learning how to use reflection in learning. *BMC Med Educ* 2007;7:5.
- 49 Rees CE, Sheard CE. Undergraduate medical students' views about a reflective portfolio assessment of their communication skills learning. *Med Educ* 2004;38(2):125-128.
- 50 Amsellem-Ouazana D, Van Pee D, Godin V. Use of portfolios as a learning and assessment tool in a surgical practical session of urology during undergraduate medical training. *Med Teach* 2006;28(4):356-359.
- 51 Overeem K, Driessen EW, Drenthe A, Van Tartwijk J, Van der Vleuten CPM. Kwaliteitscriteria voor portfoliobeoordeling van geneeskundestudenten. (in Dutch with summary in English) *Dutch J Med Educ* 2005;24(2):64-71.
- 52 Melville C, Rees M, Brookfield D, Anderson J. Portfolios for assesment of pediatric specialist registrars. *Med Educ* 2004;38:1117-1125.
- 53 Rees CE, Sheard CE. The reliability of assesment criteria for undergraduate medical students' communication skills portfolios: The Nottingham experience. *Med Educ* 2004;38(2):138-144.
- 54 McCready T. Portfolios and the assesment of competence in nursing. *Int J Nurs Stud* 2007;44(1):143-151.
- 55 Pinsky LE, Fryer-Edwards K. Diving for PERLS - Working and performance portfolios for evaluation and reflection on learning. *J Gen Int Med* 2004;19:582-587.
- 56 Wade RC, Yarbrough DB. Portfolios: A tool for reflective thinking in teacher education. *Teach Teach Educ* 1996;12(1):63-79.
- 57 McMullan M. Students' perceptions on the use of portfolios in pre-registration nursing education: A questionnaire survey. *Int J Nurs Stud* 2006;43:333-343.
- 58 Snyder J, Lippincott A, Bower D. The inherent tensions in the multiple uses of portfolios in teacher education. *Teach Educ Q* 1998;25(1):45-60.
- 59 Tigelaar DEH, Dolmans DHJM, Wolfhagen HAP, Van der Vleuten CPM. Using a conceptual framework and the opinion of portfolio experts to develop a teaching portfolio prototype. *Stud Educ Eval* 2004;30(4):305-321.
- 60 Gallagher P. An evaluation of a standards based portfolio. *Nurse Educ Today* 2001;21:409-416.
- 61 Pitts J, Coles C, Thomas P, Smith F. Enhancing reliability in portfolio assesment: Discussions between assessors. *Med Teach* 2002;24(2):197-201.

- 62 Pitts J, Coles C, Thomas P. Educational portfolios in the assessment of general practice trainers: reliability of assessors. *Med Educ* 1999;33:515-520.
- 63 Frohna JG, Kalet A, Kachur E, et al. Assessing residents' competency in care management: report of a consensus conference. *Teach Learn Med* 2004;16(1):77-84.
- 64 Lonka K, Slotte V, Halttunen M, et al. Portfolios as a learning tool in obstetrics and gynaecology undergraduate training. *Med Educ* 2001;35(12):1125-1130.
- 65 Collins A, Joseph D, Bielaszyc K. Design research: Theoretical and methodological issues. *J Learning Sci* 2004;13(1):15-42.
- 66 Mucklow JC. Continuing Medical Education in Clinical Pharmacology and Therapeutics: report of a questionnaire survey. *Br J Clin Pharmacol* 2001;52(1):9-16.

Table 1. Studies examining the use of portfolios in preclinical undergraduate medical education

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio Content; Mentoring; Assessment)	(Aim; Effect level; Findings	Limitations
Driessen et al.33 2003 To evaluate first-year students' perceptions on reflective portfolio use.	Single group, no comparison, post-test only, interviews.	38 first-year undergraduate students; 15%; Preclinical	Aim: mentoring and assessment of reflective skills; Content: reflections and materials (open structure); Mentoring: yes by staff mentor; Assessment: yes.	Level 1; 95% of students perceived portfolio as an important stimulus for reflection, 80% of students perceived reflection as useful, one of the competence areas in the portfolio proved difficult to perform in practice.	Limited number of respondents, limited information on study methods.
Driessen et al.32 2005 To examine the conditions for successful use of portfolios.	Single group, no comparison, qualitative study, grounded theory.	13 staff mentors for first-year undergraduate students; 100%; Preclinical	Aim: mentoring and assessment of reflective skills; Content: reflections and materials (open structure); Mentoring: yes by staff mentor; Assessment: yes.	Level 1; Portfolios are valuable for assessing and developing reflective skills, provided the following conditions are met: appropriate portfolio structure, appropriate assessment procedure, enough new experiences for student, sufficient teacher capacity for mentoring and assessment.	Teachers' perspective only

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Finlay et al.21 1998 To evaluate effect of portfolio learning.	Randomized controlled trial, post-test, learning outcomes, content analyses.	68 students study group and 69 control group; Study group 85% and control group 87%, content analyses 31%; Preclinical third year undergraduate longitudinal oncology program.	Aim: mentoring and monitoring of learning (understanding of the impact of malignant disease); Content: reflections, materials, overviews (diary with experiences) (open structure); Mentoring: yes, bimonthly by a trained tutor in tutorials (6-8 students); Assessment: no.	Level 2; No differences on OSCE scores, except for poor performing students (<median); they had significantly better OSCE results (P=0.04), portfolios were of a high standard.	Limited and selective response for content analyses because of voluntary submission of portfolios.
Gordon22 2003 Evaluation of student and faculty perceptions.	Single group, no comparison, post-test questionnaire.	195 first-year students and 31 faculty members; Students 67% and faculty 100%.	Aim: mentoring and assessing personal and professional development (reflective skills); Content: reflections, materials (open); Mentoring: no; Assessment: yes, by interview.	Levels 1 and 3; Students perceived the portfolio and interview as worthwhile (91%), stimulating reflection (96%) and offering opportunities to change their approach (76%); Faculty perceived the portfolio and interview positively.	Relatively low response rate among students, no pilot testing of questionnaire.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Maughan et al.34 2001 To evaluate portfolio use in an oncology program.	Single group, no-comparison, post-test questionnaire, qualitative analyses.	144 third-year students; 81%; Preclinical third year undergraduate longitudinal oncology program.	Aim: mentoring and monitoring learning (understanding of the impact of a malignant disease); Content: reflections, Materials, overviews (diary with experiences) (open structure); Mentoring: yes, monthly by a trained tutor in a tutorial group (6-8 students); Assessment: yes.	Level 1; Large majority of students found the available time appropriate for work required for portfolio; Some students had difficulty in understanding the flexibility of portfolio.	Only limited part of study focused on portfolio, no pilot testing of questionnaire.
Overeem et al.51 2005 To evaluate validity of portfolio assessment	Single group, no comparison.	Stratified sample of 40 first-year students' portfolios; 17%; Preclinical	Aim: mentoring and assessment of reflective skills; Content: reflections and materials (open structure); Mentoring: yes by staff mentor; Assessment: yes.	NR; Moderate to good inter-rater agreement (.62 to .94), the strongest predictor of the variance in the regular ratings was 'quality of reflection' (R 0.80; R2 66%), no further items accounted for a significant proportion of variance, irrelevant items, such as writing and lay-out, had negligible effects.	Limited sample.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Rees and Sheard ⁵³ 2004 To study reliability of portfolio assessment	Single group, no comparison, correlational study.	100 portfolios of second-year undergraduate students; 50%; Preclinical two-year communication skills training.	Aim: stimulate and assess reflective skills; Content: reflections and materials (open structure); Mentoring: no; Assessment: yes	NR; Moderate to high intra class correlations for total score 0.771 (95% CI=0.678, 0.840), Individual items between K0359-K0.693.	
Rees and Sheard ⁴⁹ 2004 To evaluate perception of preclinical medical students of portfolios.	Single group, no comparison, post-test only, questionnaires.	178 second-year students; 86%; Preclinical two-year communication skills training.	Aim: stimulate and assess reflective skills; Content: reflections and materials (open structure); Mentoring: no; Assessment: yes	Level 1; Most students had a neutral perception of portfolios, self-assessment of reflective skills was significantly positively related with their perception of usefulness of portfolios.	48% of respondents had no actual experience with the portfolio at the time of questionnaire completion, only total score is reported, no item scores; limited pilot testing of questionnaire.
Vink et al. ²⁹ 2005 To identify different types of reflections.	Single group, no comparison, content analyses.	18 first-year portfolios; 5%; Preclinical community-based work project.	Goal: stimulate reflection; Content: reflections (prestructured); Mentoring: yes; Assessment: yes.	Level 2 It appeared difficult to identify different types of reflection, the frequency of different types of reflection was low, except for the category evaluation.	Limited number of respondents; self selected volunteers.

Abbreviations: NA: not available; NR, not relevant

Table 2. Studies examining the use of portfolio in undergraduate clinical clerkships

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Ansellem- Ouazana et al.50 2006 To evaluate portfolio use	Single group no comparison, post-test questionnaire	24 5th an 6th year students; NA; practical training sessions of surgery.	Aim: stimulate reflective ability; Content: reflections, learning goals, materials (open structure); Mentoring: monthly meetings with a tutor; Assessment: yes, by interview.	Level 1; Students perceived portfolio as useful and acknowledge its pedagogical values.	Limited number of respondents, sampling frame unclear, no pilot testing of questionnaire.
Burch and Seggie47 2005 NA	Single group no comparison, correlational study	181 students; 100%; Clerkship	Aim: assessment; Content: patient descriptions, reflections; Mentoring: no; Assessment: interview	NR; Moderate correlations (R=0.42 with mc examination, R=0.37 with clinical case based examination), portfolio is not very labor intensive.	No clear research question, statistical results are not related to theory, no information about study methods.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Davis et al.36 2001 Evaluation of portfolio as a final examination method.	Post-test only, questionnaire, correlational study.	126 final year undergraduate students and 33 examiners; Students: 100% intervention and 83% survey, examiners: 100%; Clerkships	Aim: final examination; Content: diverse evidence material (prescribed by faculty); Mentoring: no; Assessment: yes.	Level 1; Moderate correlations (R=0.42 with mc examination; R=0.42 with essay test; R=0.47with OSCE), examiners: reading time was high burden, support for portfolio as an examination instrument; students: general support for portfolio as an examination method, too much paperwork, need for more information on how to construct the portfolio.	Research question not very specific, statistical results are not related to theory, no pilot testing of questionnaire.,
Duque et al.44 2006 To evaluate effect of hands-on introduction and the use of portfolio as assessment instrument and the effect on student-tutor interaction.	Single group, no comparison, post-test questionnaire.	133 students and 18 tutors; Students 100%, 98% survey and tutors 60%; Four-week third-year clerkship geriatric medicine.	Aim: mentoring; stimulate reflection and assessment; Content: overviews; self reflections (digital); Mentoring: yes; Assessment: yes.	Levels 1 and 2; Hands on introduction has significant positive effect on portfolio scores and comfort of students with portfolio, students valued the portfolio moderately as an evaluation tool, tutors found it an acceptable tool, but were sometimes limited by time constraints, there were many digital student tutor interactions in the portfolio (mean 30, sd 5 per student/month).	Limited response among tutors, questionnaire limited in items, no pilot testing of questionnaire.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Elango et al.37 2005 Evaluate students' perceptions of portfolio as a learning tool.	Single group, no comparison, post-test, questionnaire.	143 students in clerkships; 55%; clerkships.	Aim: assessment of reflective learning; Content: material (case reports) reflections; Mentoring: no; Assessment: yes.	Level 1; 65% of students were positive about portfolio as a useful learning tool, 93% find portfolios stressful, 67% did not regularly update portfolio, 81% find that they were not adequately mentored.	Limited response, no pilot testing of questionnaire.
Grant et al.48 2007 To evaluate the use of a reflective portfolio	Single group no comparison, post-test questionnaire	81 final year students; 96% students and 77 % marked portfolios; 3 weeks GP attachment.	Aim: mentoring and assessment of reflection; Content: learning plan, significant event analysis, health needs analyses, clinical audits; Mentoring: by GP teacher Assessment; yes	Level 1; Moderate inter-rater reliability 0.65, for raters marking only small (2 or 3 portfolios) inter-rater reliability was low 0.32 and 0.16, students valued mentoring by GP teacher, many students found the portfolio structure too prescriptive, a mayor issue for students is the large volume of written work	Limited number of respondents, sampling frame unclear, no pilot testing of questionnaire, short timeframe of the study.
Lonka et al.64 2001 To evaluate portfolio used in undergraduate training.	Single group no comparison, content analyses (quantitative and qualitative).	91 fifth-year undergraduate students; 93%; Clerkships.	Aim: coaching of learning and reflection; Content: organized per training session, overviews of clinical procedures, self assessments, Mentoring: no; Assessment: no.	Level 1; 50% of the students perceived portfolio as useful and 16 % perceived portfolio negatively (labor and time intensive), students who write more and/or are more active in the clerkship produce portfolios of higher quality.	Quantitative analyses give information that is not very relevant for research question, possible bias of qualitative evaluation, which was included in the portfolio document and therefore not anonymous.

Abbreviations: NA, not available; NR, not relevant

Table 3. Studies examining the use of portfolio in specialist training

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Fung et al.42 2000 To determine effects of portfolio on residents' perception of self-directed learning.	Non-equivalent controls, post-test questionnaires.	15 study group and 26 control group; NA; Obstetrics/gynecology residents.	Aim: monitoring learning experiences and stimulate reflective thinking; Content: overviews of patient encounters and critical incidents (electronic); Mentoring: no; Assessment: no.	Level 1; Positive effect on residents' self-perception of self-directed learning abilities.	Limited number of respondents, sampling frame unclear, only self-perceived abilities.
Kjaer et al.38 2006 To explore GP trainees perceptions of portfolio use.	Single group no comparison, post-test questionnaire	90 GP Registrars; 72%; Specialist training.	Aim: support reflective learning; Contents: learning diary, significant events descriptions, learning objectives (electronic); Mentoring: GP trainers (sometimes not familiar with the portfolio); Assessment: no.	Level 1; The portfolio helped to gain an overview and to focus on learning objectives, it had a moderate impact on the structuring and organizing their training, and stimulated reflection, lack of time, insufficient instruction and introduction of the portfolio and in some cases inappropriate IT facilities, were perceived as a problem	Limited introduction of the portfolio.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
<p>Melville et al.⁵²</p> <p>2004</p> <p>To evaluate reliability and validity of portfolio assessment.</p>	<p>Single group no comparison, correlational study.</p>	<p>105 portfolios of pediatric specialist registrars; 98%;</p> <p>Specialist training.</p>	<p>Aim: to broaden the scope of the competencies expected of registrars and help mentors to give feedback;</p> <p>Content: materials (open structured by competencies);</p> <p>Mentoring: extensive feedback after assessment;</p> <p>Assessment: formative.</p>	<p>NR;</p> <p>Feedback improves portfolio documentation, but not the standard of the assessment items. Moderate inter-rater reliabilities, correlation 0.52 and K 0.35; moderate correlations with in-training assessment 0.51 (K 0.26) and 0.25 (K 0.29) (moderate); For a G- coefficient of >0.8 4 raters are required or 1 rater on 5 occasions.</p>	<p>Portfolio was still in developmental stage, some of the outcome measures were not very informative, limited number of respondents for calculating outcome measures (N = 31 for inter-rater reliability, N = 23 for effect of feedback).</p>
<p>O'Sullivan et al.⁴⁶</p> <p>2004</p> <p>To evaluate reliability and validity of portfolio assessment.</p>	<p>Single group no comparisson, correlational study.</p>	<p>18 out of 22 (82%) portfolios of psychiatry residents, 8 out of 18 (44%) residents and 4 faculty were interviewed;</p> <p>Specialist training</p>	<p>Aim: assessment of residents competences;</p> <p>Content: evidence material (prestructured);</p> <p>Mentoring: no;</p> <p>Assessment: yes by two external raters.</p>	<p>Level 1;</p> <p>Generalizability study shows that 2 raters are required for relative decisions and 3 for absolute decisions R=0.23 with clinical rating, R=0.40 with knowledge test (modest evidence for validity), worry of residents for time, workload and fairness, portfolio gave insight into curricular weaknesses.</p>	<p>Limited number of respondents.</p>

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Pearson and Heywood35 2004 To evaluate feasibility of portfolio use for GP registrar training, identify factors that influence feasibility, consider the added value towards supervision.	Post cross-sectional survey, questionnaire, interviews and content analyses.	71 GP registrars; 77%, 7 (8%) GP registrars were interviewed and the content of their portfolios analyzed; Specialist training.	Aim: NA Content: overviews (logbook) (prestructured); Mentoring: GP supervisors (not informed about portfolio and poorly motivated); Assessment: no.	Level 1 and 2; 65 % used portfolio to record learning, 42% used portfolio for reflective learning, 33% used it to plan their learning, registrars with supportive trainers used the portfolio more in reflection, registrars in their first attachment are more likely to use the portfolio (in reflection), prescribed portfolio format was perceived to be non-adequate.	Portfolio implementation, (no clear goal, uninformed and poorly motivated mentors, structure of a logbook); no pilot testing of questionnaire.
Ryland et al.39 2006 To explore experiences of portfolio use in	Single group no comparison, post-test questionnaire	147 F2 trainees; 65%; Foundation year (F2)	Aim: monitoring of learning and training, and assessment; Content: multi-source feedback, case-based discussions, mini-CEX, observations of practical procedures, critical incidents; Mentoring: yes; Assessment: no.	Level 1; The majority of the trainees appreciate the portfolio, but understanding of purpose and content by educational supervisors is paramount to its success.	Relative low response rate, no pilot testing of questionnaire.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Snadden41 1996 To evaluate the acceptability and usefulness of portfolios for GP trainers and registrars.	Single group no comparison, action research, interviews, qualitative analyses.	20 pairs of GP registrars and trainers; 80%; Specialist training	Aim: to stimulate reflective learning; Content: reflections and overviews (logbook with critical incidents) (open structure); Mentoring: by GP trainer; Assessment: no.	Level 1; Portfolios worked well for part of the registrars, for non-reflective learning styles it is less suitable, the trainer appears to be important, trainers value portfolio more than registrars, time shortage is a potential barrier.	
Snadden and Thomas43 1998 To evaluate the usefulness of portfolio for GP vocational training.	Single group no comparison, action research over a 2-year period, interviews, focus groups, qualitative analyses.	44 GP registrars and 27 GP trainers; Registrars 100% and trainers 100%; Specialist training.	Aim: to stimulate reflective learning; Content: reflections, materials (critical incidents), overviews (reflective journal) (open structure); Mentoring: GP trainer and an external facilitator, workshops for registrars and trainers; Assessment: no.	Level 1; Portfolios were effective for reflective learning, enhanced relation trainer and registrar and provided registrars with feedback, essential factors for successful portfolio use: motivated trainers; 'concerns' with own performance on the part of the registrar; assessment.	
Webb et al.40 2006 To evaluate portfolio implementation in general surgery residency.	Single group no comparison, post-test questionnaire	40 residents; NA; general surgery residency.	Aim: stimulate reflective learning; Content: case descriptions; Mentoring: monthly by mail; Assessment: no.	Level 1; Monthly faculty review of portfolio and perceived importance by faculty were key factors for portfolio use, new residents perceive portfolio more positive than upper level residents.	Limited number of respondents, sampling frame unclear, limited information on study methods

Abbreviations: NA, not available; NR, not relevant

Table 4. Studies examining the use of portfolio in Continuing Medical Education (CME)

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
Campbell et al.30 1996 How is the portfolio used and what are the perceptions of the users of the influence on professional development?	Single group, no comparison, content analyses, post survey focus groups.	152 physicians from diverse specialties; 94% participated in survey; CME	Aim: monitoring and planning of CME, promoting reflective thinking; Content: overviews (logbook) (digital); Mentoring: no; Assessment: no.	Level 1 and level 3; Both survey (75%) and focus groups suggest that portfolio stimulates reflective learning and helps monitoring and planning of CME.	Self-selected volunteers, no pilot testing of questionnaire.
Dagley and Berrington31 2005 Evaluate feasibility and added value of portfolio for GP appraisal, revalidation and clinical governance	Pre-post test only, questionnaires, interviews, qualitative data sources, content analyses.	23 general practitioners; 7%: CME	Aim: to coach and monitor professional development of GPs; Content: overviews, material and PDPs (digital); Mentoring: no; Assessment: no	Level 1; Only 5 out of 23 (22%) GPs used the portfolio, because of time constraints and technical software problems, the content of the portfolios was rich and varied, the PDPs were of good quality, formal audits were not included in the portfolio	Small N, self-selected volunteers, short time scale for intervention, no pilot testing of questionnaire.
Dorman et al.20 2002 Evaluation of how	Single group no comparison, post-test only, questionnaire.	95 consultants in diabetes/ endocrinology; 22%;	Aim: monitoring and planning of CME, promoting reflective thinking; Content: overviews (logbook)	Level 1; Limited use by learners (34%) due to lack of time, workload, lack of IT skills, considerable support	Small N, self-selected volunteers, no pilot testing of questionnaire.

Source; Research Question	Study design	Population; Participation rate; Setting	Portfolio (Aim; Content; Mentoring; Assessment)	Effect level; Findings	Limitations
the portfolio is used and what the perceptions are of the users		CME	(digital) (open structure); Mentoring: no; Assessment: no.	for reflective learning with portfolio, but also reactions of strong dislike.	
Mathers et al.28 1999 To evaluate effectiveness and efficiency of portfolios for CME.	Quasi experimental study, pre and post survey questionnaire, interviews, observations, content analyses.	16 GPs study group; 16 GPs control group; 100%; CME.	Aim: postgraduate educational accreditation; Content: overviews, learning plans, material (critical incidents), Reflections, Mentoring: by a tutor and organized in co-mentoring groups; Assessment: yes	Levels 1 and 2; Stimulated GPs in self-directed learning and reflection, flexibility of portfolio format was highly valued, co-mentoring groups were stimulating, the combination of mentoring and assessment by mentor/facilitator was not an issue for the GPs, it was paper and time intensive and a 'slimmer' portfolio is recommended (m24.5 hours, sd 12), more deep learning goals than traditional CME.	Self-selected volunteers, no pilot testing of questionnaire.
Tate et al.45 1999 To evaluate utility of portfolio assessment for RCGP Membership examination.	Single group, no comparison.	Experienced GPs; 100 examiners; 100%; CME.	Aim: assessment for RCGP Membership examination; Content: 15 video taped consultations, logbook and consultation description forms; Mentoring: by a mentor or trainer; Assessment: yes.	Level 1; Acceptable interrater reliability: Cohen's kappa 0.78; Phi coefficient 0.80; for examiners equally time consuming as traditional oral examination, candidates find it onerous and time consuming.	Very limited information about sampling, setting, selection, and methodology.

Abbreviations: NA, not available; NR, not relevant

8 The self-critical doctor: helping students become more reflective

**Erik W Driessen
Jan van Tartwijk
Tim Dornan**

Published in: British Medical Journal 2008; 336: 827-830

Victor, a student under your supervision, can be cold and abrupt with some patients. He is a good doctor and you are sure he could be a better communicator. You have addressed this problem, but Victor clearly fails to understand what is going wrong. How do you encourage him to be more reflective?

Defining reflection

“Reflection” means letting future behaviour be guided by a systematic and critical analysis of past actions and their consequences

Why reflect?

Whether or not “experience” means “making the same mistakes with increasing confidence over an impressive number of years”¹ depends on how self analytical and critical you are. When you speak of your students “being more reflective” you mean they should let their behaviour on future occasions be guided by systematic and critical evaluation and analysis of actions and beliefs and the assumptions that underlie them.² All UK doctors are now expected to make reflection a critical foundation of their lifelong learning³ on the assumption that patients will benefit.⁴ This emphasis on reflective learning in medical education is relatively new and there is certainly no hard evidence available yet that patients benefit directly from doctors’ reflective learning.⁵ However, there is evidence that suggests reflection could help your student address his deficient communication skills. A study in postgraduate medical education found that reflection plays a vital role in residents’ learning from clinical experiences.⁶ Acquiring reflective learning skills helped undergraduate medical students identify their learning needs and stimulated learning focused on comprehension and understanding.⁷ Research in the field of expertise development,⁸ nurse training,⁹ and teacher education¹⁰ all provide evidence that reflection plays an important part in learning from experience. Students do not adopt reflective learning habits spontaneously,¹¹ so teachers must help them. The paragraphs that follow suggest how to teach reflective learning. The teaching methods are based on published studies,^{8,12} the recommendations of leading medical educators,¹³ and experience from training clinical teachers, students and residents internationally.

How can teachers foster students' reflective skills?

Prerequisites for reflection in clinical settings

To become a better communicator, your brusque student needs time to reflect and a safe, open atmosphere,¹⁴ 2 things that may be missing in the no-nonsense climate of a busy clinical workplace. You can create moments for reflection by using the time-efficient "1 minute preceptor" microskills, which provide for making a diagnosis, teaching new knowledge, and providing feedback in 4 quick steps.¹⁵ If he becomes defensive, you could make the situation safer by emphasising that it is a learning situation implying that perfection is not (yet) required and you could model reflection by describing 1 of your own communication mishaps. You should also schedule a one-to-one debriefing at a later time because good mentoring needs protected time as well as opportunistic contact.

Critical self-assessment and the identification of learning needs is fundamental to reflective learning; however, students from a traditional apprenticeship background may find it unfamiliar and threatening.¹⁰ Moreover, research shows inherent limitations in self-assessment.¹⁶ External sources of information, like practice guidelines and feedback, can enhance it^{17,18}, but students need more than self-assessment to identify their learning needs, which is where mentoring and the use of a portfolio come in.¹⁹

Portfolios

Portfolios are instruments to promote reflection. Compiled by learners themselves, they contain evidence of how goals were met and competence progressed. They contain, for instance, reports and presentations made by students themselves, feedback, assessments, and context descriptions. Often, they also include students' own written self-assessments, analyses of task performance, and plans for improving competence. They may be digital or paper based and their content may be prescribed or left to students' discretion. A recent literature review shows a flexible format to be preferable.²⁰ Too much prescribed content and too many detailed directions about how to compile and present the portfolio, can easily result the feeling that compiling a portfolio has to do more with bureaucracy than with learning and may force learners to search for content outside their direct and lived experiences.

Teaching methods

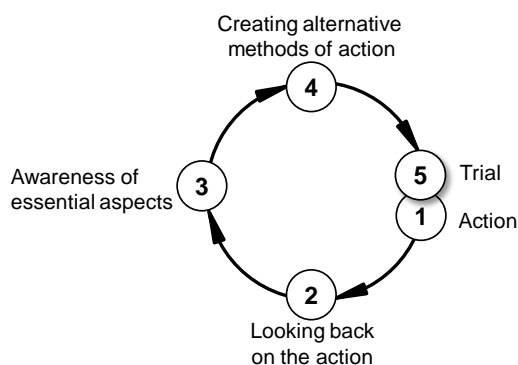
Korthagen *et al.* designed the “ALACT” model (Figure 1) to describe the cyclical process effective learners go through when faced with a situation for which no *routine* solution is available.¹⁰

Significant incident discussion, audit, peer mentoring, and portfolio can all be used to “scaffold” the ALACT model. The following paragraphs build on Korthagen *et al.*’s work¹⁰ to provide step-by-step recommendations.

Action

The reflection cycle starts with *action*. You could help your abrupt student improve his existing routines and concurrently acquire new ones by pre-selecting a mixture of clinic patients who are more or less easy to communicate with. Ericsson’s research predicts that expertise will grow not just from weight of experience, but from engaging in activities specifically designed to improve performance.²¹

Figure 1. ALACT model describing the phases of “spiral” professional development



Looking back on action: Self-assessment

Encourage the student to *look back* on informative patient encounters (positive or negative). “Looking back on action” can be regarded as *self-assessment*.¹⁹ Your student may not even know he comes across as cold and brusque until he reviews *evidence* of this in his portfolio. Student dependent *evidence* can include log books, case reports, clinical data and research projects. Student independent *evidence* comes from multi source feedback,²² mini-CEX²³ (i.e. snapshot observations of performance), direct observation of practical procedures, audits, and case-based discussions. However, feedback is of little value without critical analysis by your student.

Awareness of essential aspects: Analysis

Analysis is examining the data, seeing patterns, and identifying cause-effect associations. Your student should ask himself: What are the essential aspects of this experience? Why did things happen the way they did? During appraisals, you can help by kindly, but persistently, asking the question “why?” Pose ‘but/confronting’ questions to help him see discrepancies in his analysis. Help him see general patterns by asking questions like “Has this happened before?” Help him get a “helicopter view” by showing links between apparently discrepant pieces of information.

You review Victor’s portfolio and find this further evidence of poor communication:

A patient with chronic fatigue syndrome refused to be seen by Victor on an outpatient follow-up visit.

He scored low on a mini-CEX because he failed to explain the problem to a patient with analgesic induced headaches.

On the other hand, the portfolio shows that other patients were very positive about their encounter with this young doctor. His written self-assessment shows he is troubled by the negative evaluations, but blames them on the particular circumstances of those consultations.

In a one-to-one meeting with Victor, you contrast his warm, empathic communication with a patient who was diagnosed as having terminal cancer with the situations in which his communication skills were less effective. You stimulate him to analyse the differences between the consultations with negative evaluations and the situations in which he did well. You do this by asking questions like: Do you recognise the feedback that you received? Do you see similarities? What are the differences between the situations in which you did well and the situations in which your performance was evaluated negatively? What do you normally do when you have no explanation for a patient’s symptoms? In what ways might that come across? What did you feel when you had to deal with the problems involving uncertainties? From this discussion, Victor realises he has no strategies for dealing with uncertain situations, like patients with no clear diagnosis.

Creating or identifying alternative methods of action: Change

Following the analysis, your student must now choose *alternative methods of action*. Your role is to encourage him to consider alternatives, choose 1 of

them, and justify his choice. A SMART (Specific, Measurable, Acceptable, Realistic, Time bound) action leads into the next cycle of reflective learning.

Victor's "SMART objective" is to observe his experienced teacher in 5 consultations with patients with unexplained symptoms. He debriefs on them with his teacher and decides which communication skills he wants to develop. He applies those skills in subsequent consultations and, after 4 weeks, asks to be observed in 2 consultations.

Conclusion

Reflection underpins learning from experience, but students will not do it automatically. As a clinical teacher, your task is to stimulate students to assess and analyse their actions systematically and critically and formulate alternative actions. To do so, you must provide a challenging but safe learning environment, give feedback, and ask the right questions at the right time. As in consultations with patients, the skill of the clinical teacher is to listen well and ask open questions.

Summary points

- Reflection is vital to learning from clinical experiences, and has been shown to help learners identify their learning needs and to stimulate learning.
- Students do not adopt reflective learning habits spontaneously, so teachers must help them.
- Clinical teachers can stimulate students to assess and analyse their actions systematically and critically and formulate alternative actions.
- To do so, clinical teachers must provide a challenging but safe learning environment, give feedback, and ask the right questions at the right time.
- As in consultations with patients, the skill of the clinical teacher is to listen well and ask open questions.

Barriers to reflective learning	Supporting reflective learning
<ul style="list-style-type: none"> • Students are not used to reflection • Lack of time 	<ul style="list-style-type: none"> • Help students to structure reflection • Don't provide the answers, let them reflect themselves • Schedule one-to-one meetings • Emphasise it is a learning situation
<p><i>Action</i></p> <ul style="list-style-type: none"> • Tasks are too easy or too difficult 	<p><i>Action</i></p> <ul style="list-style-type: none"> • Help students obtain experience that supports learning
<p><i>Looking back on action</i></p> <ul style="list-style-type: none"> • Unsafe environment. Students will be reluctant to acknowledge mistakes • Lack of information 	<p><i>Looking back on action</i></p> <ul style="list-style-type: none"> • Separate performance and person: making a mistake does not mean you are a failure • Be trustworthy (honest) • Acknowledge and make success explicit • Provide feedback yourself • Use "1 minute preceptor" microskills • Stimulate students to get information from various sources and, for instance, put it in a portfolio
<p><i>Analysis</i></p> <ul style="list-style-type: none"> • Success or failure are attributed to circumstances or to others • Things are not made explicit • Experiences are regarded as incidents, patterns are not recognised 	<p><i>Analysis</i></p> <ul style="list-style-type: none"> • Focus on the student's own role in success or failure • Stimulate students to take the perspective of the other • Ask "why" questions; "confronting" questions; "generalising" questions • Point at inconsistencies in the student's analysis • Help students generalise between experiences
<p><i>Create alternative actions</i></p> <ul style="list-style-type: none"> • It's your (the teacher's) solution, not the student's • The objectives are vague or irrelevant. 	<p><i>Create alternative actions</i></p> <ul style="list-style-type: none"> • Ask the student to suggest options for change • Stimulate students to formulate their plans and check whether these plans are in line with their analysis • Help focus SMART objectives (Specific, Measurable, Acceptable, Realistic, Time bound)

Teaching tips

- ***Provide challenges, not impossible or monotonous missions.*** When giving students a task, strike a balance between what they can easily handle and what will stretch them.
- ***Give explicit attention to reflection.*** Reflection is not intuitive. Train learners to reflect by going through a routine like the 1 proposed in this article. Information supporting self-assessment can be gathered systematically in a portfolio or by audit.
- ***Emphasise students' strengths and how to take advantage of them.*** Learning what made an action successful is just as valid a product of reflection as learning from a mistake.
- ***Ask questions rather than give answers.*** To become lifelong learners, students have to learn to be independent reflectors. Asking questions is a better way of developing reflective skills than giving answers. Questions that stimulate critical thought are most helpful: What are your strengths? What needs to be improved?
- ***Stimulate concreteness:*** The teacher must help students not get lost in generalities or vagueness. Ask questions that stimulate concreteness: What did you do? What did you want to find out?
- ***Allow students to make mistakes in a safe supervisory framework.*** Being a learner means your performance is not yet perfect. Students will only reflect if they feel they can assess their own actions without having to worry that their self diagnosed failure will be used against them.

References

- 1 O'Donnell M. *A Sceptic's Medical Dictionary*. London: BMJ publications; 1997.
- 2 The origins of the concept of reflection can be traced back to Dewey, who defines reflection as that which involves active, persistent, and careful consideration of any belief or practice in light of the reasons that support it and the further consequences to which it leads. Dewey J. *How we think: A restatement of the relation between reflective thinking to the educative process*. Boston: Heath; 1933.
- 3 General Medical Council. *Revalidating doctors: ensuring standards, securing the future*. London: GMC; 2000.
- 4 General Medical Council. *Tomorrow's doctors: recommendations on undergraduate medical education*. <http://www.gmc->

uk.org/education/undergraduate/tomdoc.pdf. [Accessed September 12, 2007.]

- 5 Mamede S, Schmidt HG. The structure of reflective practice in medicine. *Med Educ* 2004;38:1302-8.
- 6 Teunissen PW, Scheele F, Scherpbier AJJA, van der Vleuten CPM, Boor K, van Diemen-Steenvoorde JAAM. How residents learn: qualitative evidence for the pivotal role of clinical activities. *Med Educ* 2007;41:763-70.
- 7 Grant A, Kinnersley P, Metcalf E, Pill R, Houston H. Students' views of reflective learning techniques: an efficacy study at a UK medical school. *Med Educ* 2006;40:379-88.
- 8 Ericsson KA, Charness N. Expert performance: its structure and acquisition. *Am Psych* 1994;49:725-46.
- 9 Jarvis P. Reflective practice and nursing. *Nurse Educ* 1992;12:174-81.
- 10 Korthagen FAJ, Kessels J, Koster B, Lagerwerf B, Wubbels T. *Linking theory and practice: The pedagogy of realistic teacher education*. Mahwah, NY: Lawrence Erlbaum Associates; 2001.
- 11 Ertmer PA, Newby TJ. The expert learner: strategic, self-regulated, and reflective. *Instructional Science* 1996;24:1-24.
- 12 Driessen EW, van Tartwijk J, Overeem K, Vermunt JD, van der Vleuten CPM. Conditions for successful reflective use of portfolios. *Med Educ* 2005;39:1230-5.
- 13 GMC, Royal College of General Practitioners. *Portfolio-based learning in general practice: Report of a working group on higher professional education: December 1993*. Occasional paper 63.
- 14 Branch WT, Paranjape A. Feedback and reflection: teaching methods for clinical settings. *Acad Med* 2002;77:1185-8.
- 15 Aagaard E, Teherani A, Irby D. Effectiveness of the One-Minute Preceptor model for diagnosing the patient and the learner: proof of concept. *Ac Med* 2004;79:42-9.
- 16 Davis DA, Mazmanian PE, Fordis M, van Harrison R, Thorpe KE, Perrier L. Accuracy of physician self-assessment compared with observed measures of competence: a systematic review. *JAMA* 2006;296:1094-1102.
- 17 Hattie J, Timperley H. The power of feedback. *Rev Educ Research* 2007; 77:81-112.
- 18 Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med* 2005;80:S46-54.

- 19 Van Tartwijk J, Driessen E, Stokking K, van der Vleuten C. Factors influencing the successful introduction of portfolios. *Q Higher Educ* 2007;13:69-79.
- 20 Driessen E, Van Tartwijk J, Van der Vleuten C, Wass V. The role of portfolios in medical education: why do they meet with mixed success? A systematic review. *Med Educ* 2007;41:1224-33.
- 21 Ericsson KA. The influence of experience and deliberate practice on the development of expert performance. In: Ericsson KA, Charness N, Feltovich PJ, Hoffman RR, eds. *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press; 2006:683-704.
- 22 Wood L, Hassel A, Whitehouse A, Bullock A, Wall D. A literature review of multi-source feedback systems within and without health services, leading to 10 tips for their successful design *Med Teach* 2006;28:185-91.
- 23 Norcini JJ, Blank LL, Duffy FD, Fortna GS. The mini-CEX: a method for assessing clinical skills. *Ann Intern Med* 2003;138:476-81.

9 Discussion

Conclusions

Research question 1: Does working with a portfolio support the development of reflective skills?

One of the starting points for this thesis is the assumption that working with a portfolio supports early development of reflective skills in undergraduate medical students. Students indicate that they are capable of producing a portfolio and that it helps them in learning how to reflect (chapter 2). The qualitative study among mentors (chapter 4) and the content analyses (chapter 6) demonstrate that a portfolio can be an effective tool to induce almost all students to reflect on their experiences. The content analysis reveals a satisfactory level of reflection, of 4.1 on a 5-point rating scale. This is an encouraging finding given the high value placed on medical doctors' reflective abilities, and the central role of reflection in workplace learning and the development of generic competencies (e.g. communication, organisational and professional competencies).

These findings are remarkable in light of disappointing levels of reflection reported by some studies inside and outside of medical education.^{1,2} Research has shown that working with a portfolio does not automatically promote the development of reflective skills. The studies in this thesis reveal a number of characteristics that determine a portfolio's effectiveness in promoting reflective skills. The value of this thesis lies predominantly in the contribution of the studies to the evidence on which educators can draw in planning a portfolio with the aim to stimulate reflective skills in medical education.³⁻⁵

There are several reasons why it does not make sense to expect that simply working with a portfolio will automatically raise students' reflective skills to the desired level. We will discuss factors that proved to contribute to the successful use of portfolios for this aim.

Mentoring

The single most decisive factor for the successful use of portfolios may well be mentoring. If no provision is made for student mentor contacts in which portfolio content is discussed and feedback given, all time and effort expended on the portfolio will probably be a waste of time. First of all, many students are initially reluctant to engage in reflection, because to most young

students the purpose of reflection is not self-evident. A teacher or mentor is needed to convince students that reflection is worthwhile. Secondly, reflection does not come naturally to most students. Mentors have a supportive role to play in the various steps of the reflective process. For example, feedback on students' self assessment reports in their portfolio gives meaning to their retrospective reflection.⁶ It was also observed by mentors that they had to point out to students which questions they should ask themselves to broaden and deepen their analyses for the portfolio. Additionally, it appears to be difficult for students to look at their actions prospectively, in other words to contemplate how they might take a different approach next time. Many of the learning objectives formulated as part of the portfolio process do not rise above the level of good intentions. Here too mentors' feedback and probing and confronting questions are crucial in arriving at relevant learning objectives.^{7,8}

Our systematic review reveals that in many cases the portfolio process does not include the indispensable aspect of adequate support and mentoring, despite evidence of its positive impact. To our surprise, we found studies of portfolios that were implemented without any form of organised support. Apart from portfolios without any mentoring, there were also examples of portfolios where mentors were uninformed, untrained or neither had nor took the time to perform their mentoring task properly.

Portfolio format

The study of the effect of an electronic versus a paper portfolio suggests ideas for making a portfolio more acceptable to mentors (chapter 6). According to the mentors in our study, an electronic portfolio is more user-friendly (less time consuming) than a paper portfolio. The same study shows that students are more enthusiastic about an electronic portfolio and spend more time preparing it. Given students' resistance to a portfolio this is by no means a trivial effect.⁹ It should be noted, however, that these positive effects are strongly dependent on adequate logistics and facilities, notably enough computers and programs that run without fail. At the first occurrence of a technological hitch, the positive effect of the electronic portfolio is in danger of being annihilated and the work on the portfolio may grind to a halt (chapter 7).

Another important point to be taken from our studies is that a portfolio must be "lean". Students and mentors alike have an aversion to massive portfolios whether on paper or screen. The review shows too many instances of portfolios being literally conceived as a huge collection of materials contained in a folder. This can be indicative of lack of clarity about the

objectives among teachers as well as students. Because of uncertainty about what is required, students are sometimes advised to produce extensive portfolios. However, there are good reasons why so many portfolio definitions refer to selective and purposeful collection of materials. Teachers and students stand to gain from portfolios that are tailored to the needs of all parties involved. So, designers should create portfolios that are fit for purpose and whose content is geared to the objective to be achieved. When summative assessment is the sole objective of a portfolio, students can be asked to include only materials related to the competencies to be assessed and to mark which parts of the portfolio have special relevance to which specific competencies (so-called captions). In that way the assessor does not have to peruse the full contents but can focus on what is relevant to the assessment.

When promoting reflection is the portfolio's objective, an open structure but with clear guidelines is preferred. The signs from our mentor study and the review are that students need to be at liberty to select topics for reflection that are meaningful to them. A study by Mansvelder-Longayroux shows that trainee teachers reflect more superficially on topics that are less relevant to their day-to-day practice.¹⁰ When portfolio content is rigidly prescribed, reflections tend to become superficial and students may even make up topics. On the other hand it is important that students, for whom a portfolio is new, are offered support and clarity as to what is expected from them. This can be achieved by organising a portfolio along the lines of professional roles or a competency profile (for example the Canmeds roles), combined with guiding questions (chapter 2). Further guidance can be offered by a well-informed mentor who introduces the portfolio and explains the objectives and how these can be attained (chapter 7). One of the insights from the mentor study is that the so-called expertise reversal effect is also seen with portfolios.¹¹ Once learners know how to do something, they need less structure. In other words, as soon as students know how to reflect, they need less support in doing so. Indeed, the aim of learning how to reflect is for students to be able to go through the different steps of the reflection model on their own (chapter 8). A portfolio that forces students who know how to reflect into a prescribed format may evoke strong resistance from students (chapter 4).

Position in the curriculum

A rigidly predefined structure can become even more counterproductive when students do not have enough new experiences on which to reflect. Reflection starts by looking back on action (chapter 8). Students should have encountered enough new events as objects of reflection (chapter 4). In the absence of new events, students reflect for the sake of reflection and may

take recourse to fantasy. In the early stages of medical school, the curriculum is mainly theoretical and students have few real practical experiences. Therefore a portfolio that is scheduled early in the curriculum should not be too ambitious.

Ever since the introduction of the portfolio, its use for assessment has been debated.¹²⁻¹⁴ In this respect, the message from this thesis is unequivocal: a portfolio that is not embedded in an assessment programme will not be taken seriously by students and teachers. The mentors conclude that assessment stands in the way of neither students' honesty in their portfolios nor of their openness in conversations with mentors. Admittedly, when a portfolio is assessed there is a danger in certain circumstances that students will use an assessment-directed strategy in composing their portfolios. This phenomenon occurs when students do not have enough experience or when the required portfolio content is not aligned with their personal experiences and interests.

Research question 2: Can a portfolio be used to assess reflective skills?

The quality of assessment

The second research question is whether a portfolio can be used to assess reflective skills. This thesis gives answers to a number of aspects of this question: validity, inter-rater reliability, acceptability and feasibility. One of the key ideas put forward in this thesis is to base assessment of portfolios on criteria from qualitative research. At the start of our project, to assess or not to assess was not the only controversial question. The value of the portfolio as an assessment instrument was under attack as well. Inter-rater reliabilities in portfolio assessment, which until then had generally been low, gave particular cause for concern.^{15,16} Educational developers responded by excessive standardisation of assessment procedures with concomitant drawbacks. Some theorists concluded that the traditional psychometric model was difficult to use in portfolio assessment.¹⁷⁻¹⁹ They pointed out that portfolio assessment is primarily concerned with qualitative information and to a lesser extent with quantitative information. This prompted us to design a holistic assessment procedure using criteria from qualitative research. Among the surprising insights to emerge from this thesis are the acceptable inter-rater reliabilities in 2 of our studies (chapters 5 and 6) and the literature review (chapter 7). As it turned out, 2 to 3 raters are sufficient to reach reliable decisions on portfolios. These findings are a strong indication that portfolio assessment may not be inherently subjective.

The validity study in this thesis supports our holistic assessment procedure: quality of reflection explains 64% of the variance in ratings and none of the

other (less relevant) criteria make a substantial contribution. Moreover, the procedure is acceptable to students and mentors and requires only limited time from teachers.

Portfolio assessment procedure format

The results of this thesis lend support to the proposed holistic assessment procedure, which is partly based on criteria from qualitative research. As with any studies on validity and reliability of assessment, extrapolation of findings is only acceptable if contexts are comparable. That is why it is important to take the results of our review into account as well, for these are based on studies in different contexts. The studies reporting on successful assessment procedures show that for effective assessment the following factors are important: holistic scoring rubrics; small groups of trained assessors; and specific rater training expertise, including benchmarking and discussion between assessors (before and intermediate).

In view of the increasing popularity of portfolios as assessment instruments, we recommend that research should address the quality of portfolio assessment in a variety of contexts.

Table 1 shows the main factors that determine the successful use of portfolios

Table 1. Summary of factors promoting portfolio success

Factor	Recommendation
Goals	Clearly introduce the goals of working with a portfolio Combine goals (learning and assessment).
Introducing the portfolio	Provide clear guidelines about the procedure, the format and the content. Be cautious for problems with information technology.
Mentoring/interaction	Provide mentoring by either teachers, trainers, supervisors or peers.
Assessment	Incorporate in the assessment procedure safeguards, like intermittent feedback cycles, student involvement of relevant resource persons (including the student), a sequential judgement procedure. Use assessment panels of 2 to 3 assessors depending on the stakes of the assessment. Train assessors. Use holistic scoring rubrics (global performance descriptors).
Portfolio format	Use a hands-on introduction with a briefing on the portfolio's purpose and the procedures. Keep the portfolio format flexible. Avoid being overly prescriptive about the portfolio content. Avoid too much paper work.
Position in the curriculum	Integrate the portfolio in other educational activities in the curriculum. Be moderately ambitious for early-undergraduate portfolio use.

Limitations

Context

A limitation of this thesis is that most of the studies were conducted within the context of year 1 of the Maastricht undergraduate medical curriculum. As with any single institution study we have to be cautious in extrapolating findings to other contexts. Our research focused on 1 particular portfolio design and its flaws in 1 particular context: year 1 of Maastricht medical school. The portfolio is exclusively aimed at stimulating reflective skills in Year 1 students to prepare them for workplace learning in which reflective skills are an indispensable asset. We have attempted to compensate for these limitations by performing an extensive literature review. This afforded a broad overview of findings concerning a broad range of portfolios in a multitude of contexts. Moreover, we present a model to stimulate reflection which can be used in any context. It is as appropriate for early undergraduate students as it is for clinical training.

The concept of reflection

Korthagen points out that reflection has been defined and operationalised in a multitude of ways and that because of this empirical studies on reflection have been few.⁸ Divergent definitions and levels of reflection in particular have given rise to problems of interpretation. The definition of reflection and the design of our reflective portfolio and our reflective teaching method lean heavily on Korthagen's theoretical work. From our perspective, the primary objective of reflection is purposeful improvement of student performance and increasing awareness in students of both the strengths and weaknesses of their performance and their ability to attribute these to external and internal causes. This definition is eminently suited to the medical education domain, where doers tend to prevail.²⁰ Our qualitative study among mentors shows that this definition is also in line with mentors' ideas (chapter 4). For the mentors, the purpose of reflection is to give students directions for improvement by identifying strengths and weaknesses in their performance. Reflection enables identification of causes that help illuminate why things are the way they are. To the mentors, professional attributes are the prime focus of reflection.

Obviously, another definition and another level of reflection would have yielded different outcomes. If we had operationalised reflection differently, for example if we had focused on reflection on students beliefs about health care and their role as a medical doctor, our portfolio design would have turned out differently.

Implications for theory and further research

The introduction of this thesis addresses 2 issues, namely the role of portfolios in medical education and the place of reflection in medical education. We will discuss the implications of this thesis on these 2 issues and also address some of the many unanswered questions that need more future research. The focus is on the implications for theory and further research, implications for practice are already discussed in the “discussion section” of this chapter and summarised in table 1.

Reflection

At the start of our project there was limited research-based knowledge on reflection within medical education. That is why we turned to 2 other domains, those of teacher education and nurse education for a definition of reflection and an educational model for learning to reflect. When we studied the literature in these domains we discovered substantial differences between professional cultures in how reflection and levels of reflection are defined and in standards of required levels of reflection. The literature of teacher education and nursing education distinguish many more levels of reflection and additionally aim for a more far-reaching type of reflection than is customary in the medical education literature.

For example, Al-Shehri categorises reflections of *general practitioners* into three levels: descriptive, analytical and evaluative.²¹ Goodman too distinguishes three levels of reflection for teachers,²² although the third level goes much deeper than Al-Shehri’s level 3. At the third level, teachers acknowledge wider issues, such as ethical and political concerns, demonstrating an understanding of how broader social forces can influence the course of their work. The model for learning from reflection (i.e. the ALACT model) includes the steps of evaluation (i.e. looking back) and analyses (i.e. Awareness of essential aspects) and has a predominant ‘technical’ objective of identifying alternative methods of action, that is more in line with the levels for reflection that Al-Sheri distinguishes than with the third level of ethical and moral concerns that Goodman describes (chapter 8). Our research gave only insight in the possibilities to teach and assess reflection. The research did not further explore the cognitive processes involved in reflection, and what its role is in the development of generic competences and learning from experiences. Research on these questions is difficult because of the different concepts of reflection and the lack of ready to use research methodologies.⁸ However, because of the importance of reflection in current medical education more fundamental research is

necessary to understand the processes involved in reflection better and how it can promote learning.

Portfolio assessment

Portfolios are seen by many as subjective and not suitable for high-stakes decisions.^{16,23} Literature learned that we need to temper our expectations for achieving strict psychometric criteria for validity and reliability, because portfolios are individualistic and non-standardised.^{15,24} It was Snadden who in a commentary for the first time asked the question if we should continue trying to fit non-standardised portfolios to psychometric criteria.²⁵ Webb *et al.* came up with the idea to use criteria derived from qualitative research instead of the psychometric criteria.²⁶ We successfully developed an assessment procedure based on this qualitative approach. Instead of a quantitative psychometric approach (i.e. looking at consistency across repeated assessment) we used a qualitative approach (i.e. adding information to the judgement process until saturation is reached). The American Accreditation Council for Graduate Medical Education (ACGME) qualified the paper (chapter 3) in which we describe our approach as: “This article represents a major shift in thinking and approaching the assessment of competence, and will enhance the usability of portfolios as an assessment tool without interfering with the goal of stimulating and supporting reflective feedback.”²⁷ The use of qualitative research criteria for portfolio assessment opens in the words of Kuper *et al.* “tantalising scope for whole programmes of research.”²⁸ These research programmes should not be limited to portfolio assessment, but also incorporate other assessments in which non-standardised qualitative information is judged, like for example work-place assessment.

Mentoring

It was, again, Snadden who already in the early eve of portfolio use emphasised that real and challenging reflection requires another person, a trainer or tutor.²⁹ We developed both a portfolio and a teaching method to organise this reflective learning. Our experiences learned that it is almost impossible to discriminate the effect of the portfolio from the effect of the teaching method (i.e. mentor discussion and feedback). Although mentoring is possible without a portfolio, it is hardly possible to use a portfolio with aim to promote reflective skills without mentoring, because mentoring is necessary for students to learn to reflect. The added value of the portfolio to mentoring is that it:

- provides the mentor insight in the student's learning and experiences;
- stimulates and also forces the student to undertake the first reflection steps (i.e. self-assessment and analyses);
- collates and integrates self-assessment with external-assessments and other information sources and gives the mentor insight in possible discrepancies, and
- it helps the student to set the agenda for mentor meetings.

We suggest more research on the interplay between mentoring and portfolio and on the competences of effective portfolio mentors.

Conditions for portfolio use

The dearth of relevant practical and theoretical information concerning a reflective portfolio for Year 1 students was reason to undertake the project reported in this thesis. The research shines new light on different practical and theoretical aspects of effective portfolio use; goals, mentoring, portfolio format, curriculum and assessment procedure (summarised in table 1). Some aspects of portfolio use are still highly controversial, especially the workload caused by portfolios.^{1,9} Future portfolio research could focus especially on the conditions related to user friendliness or feasibility of portfolios and address time constraints.

Portfolio research

In our literature review we found many studies where the description of the portfolio structure and its implementation were inadequate. In view of the wide variation in portfolio formats, researchers and peer reviewers should insist that details of portfolio structure (purpose, content, mentoring and assessment) are given, along with the context in which it was implemented, to ensure others can critically appraise their papers adequately. Although the literature indicates that portfolios are not always successful, many studies did not examine how they were implemented and why they failed. We found no studies investigating the influence of the context in which a portfolio is introduced. To claim success for an educational intervention, such as portfolios, researchers need to look carefully at the intervention in practical settings.³⁰ The implementation of a portfolio requires greater rigour than we encountered in many papers. This lack of scientific rigour may account for our disappointing finding in the review that there was no trend toward improvement of portfolio delivery over the time span represented by these studies.

Finally

We opened this thesis by referring to the significance of *generic competences*, often referred to as ‘soft skills’, for the medical profession.³¹ Reflective skills are a prerequisite for developing generic competencies, which are best acquired in a workplace environment. Students do not pick up these soft competencies automatically, they need learn how to do this. Reflection can help them acquire these competencies, and for this reason reflective skills have to be developed at an early stage of their studies. This thesis provides ample evidence that this can be done using a portfolio that meets certain conditions. We want to emphasise here what is probably the most crucial condition for an effective reflective portfolio: feedback and guidance from well-informed mentors.

Our experiences during the past 7 years when we were developing and researching our reflective portfolio has taught us that teachers’ efforts to help students acquire reflective skills early in the curriculum meet with resistance from students.³² Fortunately, the same experiences taught us also that early acquisition of reflective skills is amply rewarded later when students embark on their clinical training.

References

- 1 Pearson DJ, Heywood P. Portfolio use in general practice vocational training: a survey of GP registrars. *Med Educ* 2004;38:87-95.
- 2 Wade RC, Yarbrough DB. Portfolios: a tool for reflective thinking in teacher education? *Teaching & Teacher Education* 1996;12(1):63-79.
- 3 Pitts, J. *Portfolios, personal development and reflective practice*. ASME Understanding medical education paper. Edinburgh: ASME; 2007.
- 4 Amsellem-Ouazana D, Van Pee D, Godin V. Use of portfolios as a learning and assessment tool in a surgical practical session of urology during undergraduate medical training. *Med Teach* 2006;28(4):356-9.
- 5 Webb TP, Arahamian C, Weigelt JA, Brasel KJ. The Surgical Learning and Instructional Portfolio (SLIP) as a self-assessment educational tool demonstrating practice-based learning. *Curr Surg* 2006;63(6):444-7.
- 6 Eva KW, Regehr G. Self-assessment in the health professions: a reformulation and research agenda. *Acad Med* Oct 2005;80(10 Suppl):S46-54.
- 7 Kelchtermans, G. *Reflectief ervaringsleren voor leerkrachten (Reflective experience learning for teachers) Cahiers voor didactiek 10*. Deurne: Wolters Plantyn; 2001.

- 8 Korthagen FAJ, Kessels J, Koster B, Lagerwerf B, Wubbels T. *Linking theory and practice: The pedagogy of realistic teacher education*. Mahwah, NY: Lawrence Erlbaum Associates; 2001.
- 9 Hrisos S, Illing JC, Burford BC. Portfolio learning for foundation doctors: early feedback on its use in the clinical workplace *Med Educ* 2008;42:214-23.
- 10 Mansvelder-Longayroux DD. *The learning portfolio as a tool for stimulating reflection by student teachers*. Leiden: Leiden University 2006.
- 11 Kalyuga S, Ayres P, Chandler P, Sweller J. The expertise reversal effect. *Educ Psyc* 2003;38(1):23-31.
- 12 Snyder J, Lippincott A, Bower D. The inherent tensions in the multiple uses of portfolios in teacher education. *Teach Educ Q* 1998;25(1):45-60.
- 13 Tigelaar DEH, Dolmans DHJM, Wolfhagen HAP, Van der Vleuten CPM. Using a conceptual framework and the opinion of portfolio experts to develop a teaching portfolio prototype. *Stud Educ Eval* 2004;30(4):305-21.
- 14 Gallagher P. An evaluation of a standards based portfolio. *Nurse Educ Today* 2001;21:409-16.
- 15 Pitts J, Coles C, Thomas P. Educational portfolios in the assessment of general practice trainers: reliability of assessors. *Med Educ* 1999;33:515-20.
- 16 Koretz D, Stecher B, Klein S, McCaffrey D. The Vermont Portfolio Assessment Program: findings and implications. *Educational Measurement: Issues and Practice* 1994;13(3):5-16.
- 17 McCullan M, Endacott R, Gray MA, Jasper M, Miller CML, Scholes J, et al. Portfolios and assessment of competence: a review of the literature. *J Advanced Nurs* 2003;41(3):283-94.
- 18 Snadden D. Portfolios - attempting to measure the unmeasurable? *Med Educ* 1999;33:478-9.
- 19 Pitts J, Coles C. The challenge of non-standardised assessment of professionals – the need for a paradigm shift. *Educ Primary Care* 2003;14:397-405.
- 20 Snadden D. Portfolio-based learning and general practice vocational training. *Med Educ* 1996;30:148-52.
- 21 Al-Shehri A. Learning by reflection in general practice: a study report. *Educ General Practice* 1995;7:237-48.
- 22 Goodman J. Reflection and teacher education: a case study and theoretical analysis. *Interchange* 1984;15(3):9-26.

- 23 Roberts C. Portfolio-based assessments in medical education: are they valid and reliable for summative purposes? *Med Educ* 2002;36:899-300.
- 24 Baume D, Mantz Y, Coffey M. What is happening when we assess, and how can we use our understanding of this to improve assessment ? *Assess Eval Higher Educ* 2004;29(4):451-77.
- 25 Snadden D. Portfolios – attempting to measure the unmeasurable? *Med Educ* 1999;33:478-79.
- 26 Webb CER, Gray M, Jasper M, McMullan M, Scholes J. Evaluating portfolio assessment systems: what are the appropriate criteria? *Nurse Educ Today* 2003;23:600-9.
- 27 Accreditation Council for Graduate Medical Education *Annotated portfolio bibliography*. Washington: ACGME 2007.
- 28 Kuper A, Reeves S, Alber M, Hodges BD. Assessment: do we need to broaden our methodological horizons? *Med Educ* 2007;41(12):1121-23.
- 29 Snadden D, Thomas ML, Challis M. *AMEE Medical Education Guide No. 11 (revised): The use of portfolio-based learning in medical education*. Dundee:Association for Medical Education in Europe 1999.
- 30 Collins A, Joseph D, Bielaszyc K. Design research: Theoretical and methodological issues. *J Learning Sci* 2004;13(1):15-42.
- 31 Meng C. *Discipline-specific or academic ? Acquisition, role and value of higher education competencies*. Maastricht: Centre for education and labour market 2006.
- 32 Croonen H. Profijt van je portfolio (gain from your portfolio). *Arts in spe* 2007;November:6-9.

Summary

Chapter 1

The research reported in this thesis was inspired by recent changes in medical practice and the advent of a new teaching format, the portfolio. These two developments did not coincide accidentally. In contemporary medical practice the so-called 'soft skills', such as collaboration in teams, professionalism and communication, have come to occupy a more central position than ever before. There are several reasons why this is so: today's physicians work in large interdisciplinary teams, enter into doctor patient relationships with (often) increasingly assertive and - thanks to sources like the internet - knowledgeable patients, and, on top of that, are facing an unceasing deluge of new medical insights, whether translated into practice guidelines or not, which they are expected to incorporate into their professional practice. When things go wrong and patients complain, more often than not soft skills are involved.

In the literature and by medical educators the portfolio is being advocated as an effective method for promoting and assessing the soft skills physicians need. Helping students acquire reflective skills early in their medical training is considered especially worthwhile. When we embarked on our Ph.D. research in 2001, there was a scarcity of information from empirical studies on ways in which a portfolio could contribute to this. That is why we launched our research project aimed at providing evidence for how a portfolio can contribute to the development of reflective skills in medical students. Our central research questions were: to what extent can a portfolio contribute to the development of reflective skills and can a portfolio be used to assess reflective skills? Secondary questions were: which factors determine whether a portfolio is effective and why it is effective in some cases and not in others? Does the portfolio medium (paper or electronic) affect portfolio quality? How can a mentor promote reflective skills? Which assessment procedure is appropriate for judging the qualitative information contained in a portfolio and how reliable and valid is portfolio assessment actually?

Chapter 2

At the outset of our research we knew of no readily available portfolio design tailored to fostering and assessing reflective skills in first-year medical students. This chapter describes how we used insights from the literature in

domains outside of medical education in designing our own portfolio for this purpose. Core features of our design were: support and guidance for students from a mentor; structuring along different roles of a physician; guiding questions to stimulate reflection; and a holistic assessment procedure.

Chapter 3

Portfolios are typically not standardised. Indeed they reflect the experiences of individual students which are variegated by nature. The information contained in portfolios is predominantly qualitative, such as documentation of activities and reflections, and only a small part of it is quantitative, for instance numerical scores. However, the majority of the commonly used assessment models are based on quantitative methods, that is, on psychometrics. Psychometric tests were developed to measure stable personality characteristics, such as intelligence. Psychometrics are the bedrock of the thinking about assessment in education and its value has been amply proven in the assessment of knowledge and skills based on quantitative information. However, it falls short where assessment of complex, qualitative information is at stake: its focus on standardisation is antithetic to the central place of individual variation that characterises the portfolio. In our efforts to design a robust, holistic assessment method compatible with the non-standardised nature of the portfolio we took recourse to strategies from qualitative research: credibility and dependability in particular. Our approach yielded an elegant, sequential assessment scheme where rising doubts concerning a particular assessee prompt the collection of more judgements from more sources.

Chapter 4

We asked portfolio mentors which elements of the portfolio they considered crucial for achieving the portfolio objectives. In in-depth interviews we sought mentors' opinions about the extent to which the portfolio promotes reflection and which factors do and which factors do not make a positive contribution. Unanimously, the mentors agreed that the portfolio does indeed stimulate reflective skills. However, they also pointed out that for a portfolio to be effective certain prerequisites must be met: students must have enough experiences on which to reflect, there must be an adequate assessment method, enough dedicated time for teachers to guide students and an appropriate portfolio structure. The mentors did not think it self evident that a portfolio could be effective in itself. Whenever one of the conditions is not met, a double threat is looming: few or no learning effects and resistance to the portfolio from both students and teachers.

Chapter 5

These findings prompted us to address the question of what is actually being assessed by a portfolio. In other words, how valid is portfolio assessment? Our portfolio was specifically designed to assess the quality of students' reflections. Both in the literature and in our practical experience we found grounds for concern as to whether a holistic assessment procedure might not give assessors too much freedom and thereby introduce a risk of bias from irrelevant criteria. By irrelevant criteria we mean aspects like writing style, spelling, lay-out and presentation. At the time of the study we had at our disposal no empirical data about the validity of portfolio assessment.

We developed an instrument to analyse which criteria were used in portfolio assessment and which of these carried the most weight. Using this instrument, two researchers analysed medical students' portfolios. The regression analysis showed no significant correlation between assessment outcomes and irrelevant criteria. As we had hoped, differences in outcomes turned out to be primarily associated with the quality of reflection.

Chapter 6

The instrument developed for the study described in chapter 5 enabled us to tackle a question that has been waiting to be answered since the portfolio was first introduced in education: does it make a difference whether the portfolio is a paper one or an electronic one? In other words, does medium matter? The electronic portfolio has numerous advocates who are firmly convinced of its added value. Unfortunately, there is no empirical evidence to back up their enthusiasm and in the literature concern is voiced that reflections elicited by electronic portfolios may be more superficial.

We examined this issue in an experimental study comparing 45 electronic and 47 paper versions of an otherwise identical portfolio. We looked at quality of reflection, ease of use and student motivation. The multi-level analysis revealed no significant association between medium and quality, except for individualised content. When students produced an electronic portfolio, both form and content of the portfolio showed significantly more markedly individual characteristics. Moreover, students spent considerably more time preparing their portfolios if it was done electronically. The mentors appreciated the electronic portfolio as being more user friendly, because of the ease of searching in the portfolio and access from different locations.

Chapter 7

The research reported in this thesis shows that, provided certain critical conditions are met, the portfolio can be an effective instrument in stimulating and assessing the reflective skills of students. In chapter 7 we describe a study aimed at identifying determinants of the effectiveness of portfolios. To this end we shifted our focus from the portfolio we had designed for our first-year students to the general literature and conducted a systematic review of empirical portfolio studies to analyse the available evidence for what makes portfolios work. The results confirmed the findings reported in chapter 4. Whether a portfolio is successful or not depends on good guidance for students (mentoring), sufficient and varied experiences for students to reflect on, a clearly defined structure and an appropriate assessment method. On top of this our review yielded some surprising insights. The most striking finding was the reported good interrater reliabilities of portfolio assessment. This is diametrically opposed to findings in other domains and to the prevailing view that portfolio assessment by definition is subjective and therefore unreliable. Another striking finding is the apparent feasibility of combining guidance and assessment of portfolios. Finally, portfolio success was found to depend strongly on the amount of time and effort required.

Chapter 8

The main theme running through this thesis is the crucial role played by mentors. Earlier studies have highlighted that mentoring and portfolio are closely interwoven. Based on our growing realisation that good mentoring is the cornerstone of portfolio success we devoted our final study to the development of a method to stimulate reflection. As we had done when designing our reflective portfolio, we again borrowed insights from outside of medical education, i.e. nurse and teacher education. The main components of the method we developed are: creating a safe and at the same time challenging learning environment, providing feedback, and asking the right questions at the right moment.

Chapter 9

In the final chapter we revisit the theories and questions that were the driving force behind our research project. Our studies confirmed the theory-based notion that the portfolio is a good educational tool for stimulating and assessing medical students' reflective skills. However, this finding has to be qualified by our discovery that portfolio success is conditional on the fulfilment of certain prerequisites. Our results also confirm findings from

other studies in different contexts: the importance of good mentoring and sufficient experiences as a basis for worthwhile reflection. There were also findings that contradicted or supplemented the theory set out in the introduction, the most striking ones being the unexpected acceptable interrater reliabilities of portfolio assessment and the feasibility of combining in one portfolio the functions of promoting and assessing reflective skills.

In summary, the main implications of our research for educational practice are the following:

- make sure the portfolio is properly introduced to students and teachers alike; clearly communicate the goals and procedures;
- ensure that the portfolio is supported by mentoring. Students appreciate and learn from conversations with mentors who give feedback on the portfolio and help them set new, realistic goals and find ways to achieve them;
- ensure that the portfolio is firmly embedded within the context and procedures of the curriculum of which they are a part;
- ensure that the portfolio is assessed. Unassessed portfolios tend to fall victim to competition from educational activities that are incorporated in the summative assessment programme;
- support students by giving clear guidelines about what they are expected to do but do not curtail their freedom by overregulating the portfolio;
- take steps to make the portfolio feasible and user friendly for students and teachers alike. Put limits on the amount of time required. Prevent uninformative, cumbersome piles of paper. With electronic portfolios, ensure that the system is working smoothly and designed for ease of use by all concerned, and
- incorporate in the assessment procedure safeguards, like intermittent feedback cycles, involvement of relevant resource persons (including the student), and a sequential judgement procedure.

Fundamental research is needed to develop a greater understanding of reflection and its role in learning. We also recommend studies to examine the interaction between mentoring and portfolio and competencies for effective mentors. Finally, our successful application of criteria from qualitative research in developing an assessment method appears to open up promising new avenues for educational development and research.

Samenvatting

Hoofdstuk 1:

Recente veranderingen in het artseneroep en de opkomst van een nieuwe onderwijsmethode –het portfolio- in de artsopleiding vormden de aanleiding tot het onderzoek waarvan dit proefschrift verslag doet. Dat beide ontwikkelingen samenvallen, is geen toeval. Er wordt in de huidige medische praktijk meer dan ooit een beroep gedaan op de zogenaamde ‘zachte vaardigheden’ van de arts, zoals samenwerken, professionaliteit en communicatie. De reden hiervoor is dat de hedendaagse arts in grotere interdisciplinaire teams functioneert, behandelrelaties heeft met (vaak) mondige en - via internet - goed geïnformeerde patiënten en bovendien een steeds sneller aanzwellende stroom van al/of niet in richtlijnen vertaalde medische inzichten moet toepassen in de praktijk. Als er iets mis gaat, bijvoorbeeld bij klachten van patiënten, zijn het vaak juist de zachte vaardigheden die in het geding zijn.

In de literatuur en in het medisch onderwijs wordt het portfolio geafficheerd als een effectieve methode om de zachte vaardigheden aan te leren en te beoordelen. Met name het bevorderen en beoordelen van reflectievaardigheden in een vroeg stadium van de opleiding wordt als belangrijk beschouwd. Hoe het portfolio hiervoor succesvol kon worden ingezet, daarover waren in 2001, bij de start van dit proefschrift, maar weinig empirische gegevens beschikbaar. Wij beoogden dan ook met ons onderzoek evidentie te leveren voor de mogelijkheden en onmogelijkheden van het portfolio als leermiddel voor reflectievaardigheden. Onze centrale onderzoeksvragen waren: in hoeverre ondersteunt het werken met een portfolio de ontwikkeling van reflectievaardigheden en kan het portfolio ook worden ingezet om reflectievaardigheden te beoordelen? Subvragen waren: welke factoren bepalen of het portfolio doeltreffend is en waarom dit soms wel en soms niet het geval is? Is het medium van het portfolio van invloed op de kwaliteit? Hoe kan de mentor reflectievaardigheid verhogen? Welke beoordelingsprocedure doet recht aan de kwalitatieve informatie in het portfolio en hoe betrouwbaar en valide is deze beoordeling dan?

Hoofdstuk 2

Bij de start van het proefschrift was er geen portfolio-ontwerp voorhanden dat specifiek gericht was op het stimuleren en beoordelen van

reflectievaardigheden bij eerstejaars studenten. Hoofdstuk 2 beschrijft hoe wij bij het ontwerpen van een dergelijk portfolio inzichten hebben toegepast uit de literatuur buiten het medisch onderwijs. Belangrijke kenmerken van ons ontwerp zijn: ondersteuning en begeleiding van de student door een mentor, structurering van het portfolio aan de hand van verschillende rollen van de arts, richtvragen om reflectie te stimuleren en een holistische beoordelingsprocedure.

Hoofdstuk 3

Portfolio's worden gekenmerkt door afwezigheid van standaardisatie: ze weerspiegelen immers de ervaringen van individuele studenten, die per definitie sterk uiteenlopen. Portfolio's bevatten vooral kwalitatieve informatie, zoals beschrijvingen en werkstukken, en in veel mindere mate kwantitatieve informatie, zoals numerieke scores. De gangbare beoordelingsmodellen zijn echter gebaseerd op de methodologie van het beoordelen van kwantitatieve informatie, ook wel psychometrie genoemd. De psychometrie ligt ten grondslag aan het beoordelen van stabiele persoonlijkheidskenmerken, zoals intelligentie, en zij vormt de basis voor het denken over toetsing in het onderwijs. Voor het beoordelen van kennis en vaardigheden op basis van kwantitatieve informatie heeft de psychometrie haar waarde ruimschoots bewezen. Maar als het gaat om het beoordelen van complexere, kwalitatieve informatie schiet de psychometrie tekort: de door de psychometrie beoogde standaardisering staat haaks op het basisidee van het portfolio, dat juist uitgaat van individuele verschillen. Om een robuuste, holistische beoordelingsmethode te ontwerpen, die recht doet aan het niet gestandaardiseerde karakter van het portfolio, gingen we uit van kwalitatieve onderzoeksstrategieën met als belangrijkste: geloofwaardigheid (credibility) en betrouwbaarheid (dependability). Dit resulteerde in een elegante, sequentiële beoordelingsmethode waarbij steeds meer oordelen worden verzameld naarmate de twijfel over een oordeel toeneemt.

Hoofdstuk 4

We hebben de mentoren gevraagd welke elementen van het portfolio naar hun mening bepalend zijn voor het realiseren van de beoogde doelen. In diepte-interviews werd aan de mentoren gevraagd in hoeverre het portfolio aanzet tot reflectie en welke factoren hiertoe wel of geen bijdrage leveren. De mentoren waren unaniem van mening dat het portfolio reflectievaardigheden stimuleert. Als belangrijke voorwaarden hiervoor noemden zij: voldoende ervaringen waarover studenten kunnen reflecteren, een adequate beoordelingsmethode, voldoende docenttijd en een adequate structuur. In de

ogen van de mentoren is het niet vanzelfsprekend dat een portfolio goed werkt. Zodra aan een van de essentiële factoren niet is voldaan, ontstaat het dubbele gevaar van weinig of geen nuttige effecten en een grote weerstand tegen het portfolio bij zowel studenten als docenten.

Hoofdstuk 5

De onderzoeksvraag diende zich aan wat er nu eigenlijk met het portfolio beoordeeld wordt; met andere woorden, wat is de validiteit van de beoordeling? We beogen met het portfolio de kwaliteit van reflectie te meten. Zowel in de literatuur als in onze eigen praktijk heerste bezorgdheid dat beoordelaars bij een holistische beoordelingsmethode te veel ruimte zouden krijgen om irrelevante criteria mee te laten wegen in hun eindoordeel. Denk hierbij aan zaken als schrijfstijl en spelling of vormgeving en presentatie. Empirische gegevens over de validiteit van het beoordelen met portfolio's ontbraken op dat moment.

We ontwikkelden een instrument om te analyseren welke criteria gebruikt werden bij het beoordelen van portfolio's en welke het meeste gewicht in de schaal legden. Met dit instrument analyseerden twee onderzoekers de kwaliteit van portfolio's. De regressieanalyse gaf geen significante correlatie te zien tussen beoordeling en irrelevante criteria. Verschillen in beoordeling bleken voor een belangrijk deel samen te hangen met de kwaliteit van de reflectie.

Hoofdstuk 6

Het in studie 5 ontwikkelde onderzoeksinstrument stelde ons in staat om een onderzoeksvraag te bestuderen die al vanaf de invoering van het portfolio in het onderwijs speelt: in hoeverre maakt het verschil of studenten een papieren of een elektronisch portfolio samenstellen? Met andere woorden, beïnvloedt het gebruikte medium de kwaliteit van het portfolio? Het elektronische portfolio kan rekenen op een grote groep pleitbezorgers die meerwaarde voor deze vorm van portfolio zien. Enige empirische evidentie voor deze 'superieure' eigenschappen is echter niet voorhanden en in de literatuur spreken anderen bezorgdheid uit dat een elektronisch portfolio tot oppervlakkiger reflecties zal leiden.

In een experimenteel onderzoek hebben wij 45 elektronische en 47 papieren versies van hetzelfde portfolio vergeleken op kwaliteit, gebruikersvriendelijkheid en motivatie van studenten. Uit de multi-levelanalyse bleek dat er geen significante relatie was tussen medium en kwaliteit van het portfolio, behalve voor de persoonlijke invulling van het

portfolio. De elektronische portfolio's waren zowel qua vorm als qua inhoud significant persoonlijker. Ook werd door de studenten aanzienlijk meer tijd besteed aan het samenstellen van het elektronisch portfolio. De mentoren vonden het elektronisch portfolio gebruikersvriendelijker, door het zoekgemak en de toegankelijkheid vanaf verschillende plaatsen.

Hoofdstuk 7

De studies in dit proefschrift laten zien dat het portfolio, onder zekere voorwaarden, een effectief instrument is om reflectievaardigheid niet alleen te stimuleren maar ook te beoordelen. In dit hoofdstuk beschrijven we een onderzoek waarin gezocht werd naar empirisch onderbouwde factoren die bepalend zijn voor de effectiviteit van portfolio's. Hiertoe verlegden we onze aandacht van het door onszelf ontworpen portfolio naar algemene gegevens uit de literatuur. In een systematisch review van empirisch onderzoek naar portfolio's in het medisch onderwijs analyseerden wij de evidentie voor de effectiviteit van portfolio's. De resultaten bevestigden onze bevindingen uit hoofdstuk 4. Het al of niet slagen van een portfolio hangt af van: goede begeleiding, voldoende verschillende ervaringen om over te reflecteren, een heldere structurering en een goede beoordelingsprocedure. Daarnaast bood het review een aantal verrassende nieuwe inzichten. Wel de opvallendste hiervan is de gerapporteerde interbeoordelaarsbetrouwbaarheid. Deze is lijnrecht in tegenspraak met bevindingen in andere domeinen en met het hardnekkige idee dat portfoliobeoordelingen per definitie subjectief en dus onbetrouwbaar zijn. Nog een opvallende bevinding is dat begeleiding en beoordeling goed te combineren zijn. Als laatste vonden we dat het succes van een portfolio sterk bepaald wordt door de benodigde tijd en de werkbelasting.

Hoofdstuk 8

Als een rode draad door dit proefschrift loopt de cruciale bijdrage van begeleiding door een mentor. De eerdere studies onderstreepten dat begeleiding en portfolio onlosmakelijk met elkaar verbonden zijn. Vanuit het groeiende besef dat goede begeleiding een *conditio sine qua non* is voor het welslagen van een portfolio hebben wij het laatste onderzoek in dit proefschrift gewijd aan de verdere ontwikkeling van een methode om reflectie te stimuleren. Evenals bij het ontwerpen van het portfolio hebben wij ook hierbij inzichten van buiten het medisch onderwijs (leraren- en verpleegkundige opleidingen) vertaald naar het medisch onderwijs. De belangrijkste elementen van de methode zijn het creëren van een veilige en

tegelijk uitdagende leeromgeving, het geven van feedback en het stellen van de juiste vragen op het juiste moment.

Hoofdstuk 9

In het laatste hoofdstuk keren we terug naar de theorieën en de vragen die het uitgangspunt waren voor dit promotieonderzoek. Onze studies bevestigen het theoretische uitgangspunt dat het portfolio een goed instrument is om reflectievaardigheden te stimuleren en te beoordelen. Ze laten echter ook zien dat dit niet vanzelf gaat en dat het portfolio alleen doeltreffend is als aan bepaalde voorwaarden wordt voldaan. Onze resultaten bevestigen ook wat reeds uit onderzoek in andere contexten bekend was, het belang van goede begeleiding en voldoende ervaringen als bron voor reflectie. Er zijn ook bevindingen uit het proefschrift die theorie uit de inleiding weerleggen of aanvullen. Hiervan zijn wel de opvallendste de gevonden acceptabele interbeoordelaarsbetrouwbaarheid en het inzicht dat het portfolio een combinatie van stimulering en beoordeling van reflectievaardigheden toelaat. Kort samengevat zijn de belangrijkste implicaties van onze bevindingen voor de onderwijspraktijk:

- Het is belangrijk dat doel en werkwijze van het portfolio steeds duidelijk worden uitgelegd niet alleen aan de studenten maar ook aan de docenten die ermee moeten werken.
- Het portfolio moet gecombineerd worden met begeleiding door een mentor. Studenten waarderen gesprekken met een mentor van wie zij feedback krijgen over hun portfolio en die hen helpt nieuwe, haalbare leerdoelen vast te stellen en aanwijzingen geeft hoe die te realiseren.
- Het portfolio moet structureel ingebed zijn in de inhoud en de organisatie van het curriculum.
- Het portfolio moet opgenomen zijn in het toetsprogramma om te voorkomen dat de studenten hun aandacht verplaatsen van het portfolio naar andere, wel getoetste onderwijsonderdelen.
- Studenten moeten duidelijke richtlijnen krijgen over wat er van hen wordt verwacht zonder dat hun vrijheid te zeer wordt ingeperkt door overdreven strakke voorschriften.
- Het portfolio moet gebruikersvriendelijk zijn en geen overbelasting vormen voor studenten en docenten. De tijdbelasting moet redelijk zijn. Er moet gewaakt worden voor het genereren van nietszeggende bergen papier. Bij een elektronisch portfolio moet het systeem goed

functioneren en voor alle betrokkenen toegankelijk en makkelijk te gebruiken zijn.

- Om de kwaliteit van de beoordelingsprocedure te waarborgen kan gebruik worden gemaakt van periodieke feedback, inbreng van verschillende personen als informatiebron, waaronder de student zelf, en een sequentiële beoordelingsprocedure.

Fundamenteel onderzoek zal meer inzicht moeten bieden in reflectie en de rol van reflectie bij het leren. Ook zijn wij van mening dat onderzoek nodig is naar de interactie tussen mentor en portfolio en de competenties van een effectieve mentor. Ten slotte opent de toepassing van criteria voor kwalitatief onderzoek bij beoordelingsprocedures een scala aan nieuwe mogelijkheden voor ontwikkeling en onderzoek van onderwijs.

Curriculum Vitae

Erik Driessen was born in Wijchen, the Netherlands on 28 October 1966. After completing the bachelor program in human resource management at the College of Social Studies, Kampen and Zwolle, he received a master's degree in educational science from the University of Amsterdam in 1994. In that same year he started his career at Maastricht University as a researcher in assessment at the Faculties of Economics, Health Sciences, and Law. In 1997 and 1998 he was a student advisor and assessment coordinator at the Faculty of Law, where he remained as an assistant professor and assessment coordinator until 2000.

At the end of 2000 he was given the opportunity to coordinate the development and implementation of a portfolio to stimulate and assess self-reflection in the Maastricht undergraduate medical curriculum. This project was the basis of this doctoral thesis.

Being part of a highly motivating and inspiring research group in Maastricht with many international contacts has been a highly enriching experience. In 2007 he was a visiting professor at Gifu Medical School, Gifu, Japan.

His current research interests are:

- portfolios for learning and assessment;
- learning and assessment in the workplace, with emphasis on postgraduate specialist training;
- the self-critical doctor; how to stimulate and assess a self-critical attitude in doctors.

He is at present also involved in the nation-wide modernisation of postgraduate specialist training in the Netherlands.