



Developing

TEACHER

ASSESSMENT

John Gardner, Wynne Harlen, Louise Hayward  
and Gordon Stobart,  
with Martin Montgomery

# **Developing Teacher Assessment**



# **Developing Teacher Assessment**

*John Gardner, Wynne Harlen,  
Louise Hayward and Gordon Stobart,  
with Martin Montgomery*



Open University Press

Open University Press  
McGraw-Hill Education  
McGraw-Hill House  
Shoppenhangers Road  
Maidenhead  
Berkshire  
England  
SL6 2QL

email: [enquiries@openup.co.uk](mailto:enquiries@openup.co.uk)  
world wide web: [www.openup.co.uk](http://www.openup.co.uk)

and Two Penn Plaza, New York, NY 10121-2289, USA

First published 2010

Copyright © John Gardner, Wynne Harlen, Louise Hayward and Gordon Stobart  
2010

All rights reserved. Except for the quotation of short passages for the purposes of criticism and review, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publisher or a licence from the Copyright Licensing Agency Limited. Details of such licences (for reprographic reproduction) may be obtained from the Copyright Licensing Agency Ltd of Saffron House, 6–10 Kirby Street, London, EC1N 8TS.

A catalogue record of this book is available from the British Library

ISBN-13: 978 0 335 23783 8 (pb) 978 0 335 23782 1 (hb)

ISBN-10: 0335237835 (pb) 0335237827 (hb)

Library of Congress Cataloging-in-Publication Data

CIP data has been applied for

Fictitious names of companies, products, people, characters and/or data that may be used herein (in case studies or in examples) are not intended to represent any real individual, company, product or event.

Typeset by Aptara Inc., India  
Printed in the UK by Bell and Bain Ltd., Glasgow

**The McGraw-Hill Companies**

# Contents

|   |     |
|---|-----|
| <i>Acknowledgements</i>   | vii |
| <b>Developing teacher assessment: an introduction</b><br><i>John Gardner</i>        | 1   |
| <b>Part I: Lighting the fuse</b>  |     |
| <b>1 Assessment to support learning</b><br><i>Wynne Harlen and John Gardner</i>     | 15  |
| <b>2 What is quality teacher assessment?</b><br><i>Wynne Harlen</i>                 | 29  |
| <b>3 What is happening in the UK?</b><br><i>Martin Montgomery</i>                   | 53  |
| <b>Part II: Spreading the word, widening the practice</b>                           |     |
| <b>4 What is innovative about teacher assessment?</b><br><i>John Gardner</i>        | 71  |
| <b>5 Moving beyond the classroom</b><br><i>Louise Hayward</i>                       | 85  |
| <b>6 Professional learning to support teacher assessment</b><br><i>Wynne Harlen</i> | 100 |
| <b>7 Teachers as self-agents of change</b><br><i>John Gardner</i>                   | 130 |

## **Part III: Keeping it going**

|   |     |
|---|-----|
| <b>8 Making a difference: evaluating the impact of innovations in assessment</b><br><i>Gordon Stobart</i>   | 141 |
| <b>9 Embedding and sustaining developments in teacher assessment</b><br><i>Wynne Harlen and Louise Hayward</i>  | 155 |
| <b>Appendix 1: Projects examined and members of the core group and advisory committee of the Nuffield Foundation-funded Analysis and Review of Innovations in Assessment (ARIA) Study, 2006/8</b> | 172 |
| <b>Appendix 2: A professional development template</b><br><i>John Gardner</i>   | 174 |
| <i>References</i>   | 179 |
| <i>Index</i>  | 195 |

# Acknowledgements

In the course of our work for this book, over 200 experts from across the UK and every part of the educational community gave their time and expertise to refine our understanding of introducing change into schools across the UK. They included teachers, head teachers, researchers, local authority and government officials, inspectors, union representatives, consultants and representatives from statutory and non-statutory bodies. A smaller but still significant number of people, who were or are involved in major assessment projects, also contributed their experiences of trying to effect change through developing teacher assessment in schools. Too legion to name, we owe all of them our sincerest thanks and appreciation.

The Analysis and Review of Innovations in Assessment (ARIA) project, which has provided the empirical strength for this book, could not have proceeded so successfully without the considerable support of our Advisory Committee: Paul Black, Richard Daugherty, Kathryn Ecclestone, Mary James, Dorothy Kavanagh, Alison Kidd, Stuart Merry, Paul Newton, Catrin Roberts, Mike Walker, Anne Whipp and Martin Montgomery (who also contributed Chapter 3). A special, heartfelt 'thank you' goes to them and to both Despina Galanouli and Jo Wilson for their excellent research support.

Finally, we would very much like to thank the Nuffield Foundation for their generous support of the fieldwork and Fiona Richman of the Open University Press/McGraw-Hill for encouraging us to follow through with our ideas in the form of this book.

John Gardner, Wynne Harlen, Louise Hayward and Gordon Stobart  
June 2009





# **Developing teacher assessment: an introduction**

*John Gardner*

## **Initiative overload**

For many years now, schools and colleges across the UK have been swamped by 'must do' initiatives designed to improve learning, raise standards and meet a whole host of other complex educational aspirations. The shelves of head teachers' offices overflow with glossy guidance and resource packs – some excellent, some questionable and some downright faddish. While the tally of initiative-related materials from government bodies and local authorities competes with the commercial sector's unsolicited freebies, 'initiative overload' deprives teachers of time for creativity and development. And head teachers everywhere seek shelter from what Coughlan (BBC News, 2007a) suggests is either a 10-year 'blizzard of gimmicks' or a 'golden age'. Prominent among this onslaught of innovations, however, has been the positive emergence of research-proven formative assessment or assessment for learning (AfL) as a major ally in the pursuit of improved learning and raised standards. Its benefits have been proven in many contexts since the seminal review of Black and Wiliam (1998a) and Hargreaves (2005: 9) explains its wildfire success in schools as being due to '... the scientific evidence and the practice evidence [being] aligned and mutually supportive'.

## **Teacher assessment**

Thanks to AfL, assessment by teachers has also taken on a new importance in schools and education systems around the world, as a complement or replacement for external testing. However, there remains a public concern that assessment by teachers may be untrustworthy or inconsistent between schools. There is evidence that teachers can inadvertently or otherwise undertake assessments with a degree of bias; for example, in favouring (or in some cases putting down) their own pupils.<sup>1</sup> And consistency between schools is an enduring challenge for any assessment system. But

## 2 DEVELOPING TEACHER ASSESSMENT

there is also considerable evidence that appropriate training and effective moderation processes can make teachers' judgements as sound (i.e. reliable and valid) or even more dependable than the results from any external test. And we do not need to look further than the trusted role that teachers play in many state examinations such as in first-marking their pupils' coursework for the General Certificate in Secondary Education (GCSE). Instances of plagiarism, inappropriate levels of assistance and other problems have resulted in home-based coursework being phased out, to be replaced by 'controlled' assessment approaches from 2009 onwards (QCA, 2008a). These new arrangements for coursework include fixed time periods and supervision, with the tasks defined by the examining bodies. Crucially, it was not teacher judgement that caused the change (it was more the plagiarism threats of the Internet, etc.) and teacher assessment is retained as the first-marking process followed by moderation. With some six million GCSE entries every year the teachers' role remains vital to the successful processing of this huge enterprise.

That said, GCSEs and their ilk are externally created, their marking schemes are decided without reference to the large majority of teachers and their pupils, the test-taking process is highly regulated and formal, and the bulk of their marking is carried out by examiners commissioned by an examinations body. Teacher assessment, that is assessment by teachers, is distinctly different and is experiencing a renaissance (see Chapter 3) throughout the UK. Teacher assessment is classroom-based. It is not a school-based variant of external testing where teachers design and produce tests to be taken by the school's pupils, under the auspices of the school and marked and judged by the school's staff. At present few schools would have teachers with the necessary technical skills to do this though this situation is changing as the UK Chartered Institute of Educational Assessors promotes training in all types of assessment expertise, including AfL, in UK schools. So what is distinctive about this book's approach to assessment by teachers?

The answer is enshrined in the first of 10 principles<sup>2</sup> we offer throughout this book, all of which are designed to guide good practice. This is the principle that *assessment of any kind should ultimately improve learning*.

As the central underpinning philosophy in good assessment practice, this is a deceptively understated principle. If our ultimate goal is to help young people to prepare and be ready for a life of learning, then arguably everything we do in education should assist that goal. Assessment processes should therefore help pupils to improve their learning as well as fulfilling other roles such as helping them to gain certification, measuring the extent of their achievements and reporting these to others. Yet it remains a distant prospect that students will receive feedback on their final university examinations, rather than simply have a line on a parchment.

Even the relatively recent development of returning GCSE scripts is primarily used to look at the quality of the marking to see if the school should appeal, rather than feed back into learning.

### **Why is interest in teacher assessment growing?**

There are several key influences acting to generate greater interest in assessment by teachers, not least the soaring costs of external testing, the improbability of external tests contributing effectively to learning and the potential of such testing to distort the focus of learning. This is not a new phenomenon; indeed, the potential for examinations to miss the point of 'education' was a problem alluded to by Joshua Fitch in his 1880 lectures at the University of Cambridge on teaching:

The whole subject of examinations looms very large in the vision of the public and is apt to be seen out of its true proportions, mainly because it is the one portion of school business which is recorded in newspapers. We shall perhaps arrive at right notions about it more readily, if we first consider the business of examinations wholly subordinate to that of education, as part of the work of a school.

(Fitch, 1898: 158–9)

In more recent times in the UK, external examinations have become unpopularly associated with evaluating school performance and the performance of teachers. Many support the use of external tests in high stakes situations (e.g., admission to university) but each country in the UK has at some time experienced repeated testing of pupils in both the primary and secondary sectors. This testing, though crudely claimed to provide a standards-raising challenge to pupils, has been designed primarily to enable education authorities and parents to judge how well a school is performing. One of the most invidious outcomes of such a strategy has been the ranking of schools in 'league tables' based on the pupils' performances in externally administered tests. England alone retains this use of external testing at 11 and 16 years old, despite the tests at the end of Key Stage 3 (approximately 14-year-olds) being scrapped owing to what the Schools' Secretary for England, Ed Balls, described as a marking 'shambles' (BBC News, 2008). The massed ranks of educationalists in all sectors remain vociferously opposed to the league tables though attitudes to the national curriculum assessments are more divided. There are those who oppose them on educational grounds and those who prefer them to increased use of teacher assessment, which they perceive would cause an increase in workload (BBC News, 2009).

#### 4 DEVELOPING TEACHER ASSESSMENT

The dominance of external testing in government policy in England has created a situation in which schools are forced continuously to improve their pupils' performance to meet externally imposed targets and standards (though clearly this could also be the case if teacher assessment was the established assessment regime). Under the worthy aspiration to raise standards, governments have wittingly or otherwise caused the indicators of their standards (e.g., national curriculum test scores for primary school reading in England) to become the objectives of the pupils' learning. One of the most damaging consequences of this is that schools may begin teaching to the tests to improve performance and meet or exceed the standards. Naturally, such a strategy must weaken attention to the wider curriculum and its learning goals. This distortion in a school's curriculum and mission must also give rise to poor motivation for pupils who do not improve sufficiently or indeed consistently fail to do well over the multiple testing occasions involved.

These problems have long been recognized by educationalists and there is significant evidence pointing to the use of assessment in support of learning as being a major and more appropriate means of assisting individual pupils to reach their potential. This does not mean that external tests should be rejected; rather, it is part of an argument that calls for the rejection of their overuse, and the educational problems to which such overuse can give rise. External testing will always be a feature of most education systems and that is how it should be. However, the complementary and currently underdeveloped role of assessment by teachers holds the promise of assessment more clearly integrated into the classroom and learning context. And this is the focus of this book – harnessing the benefits of teacher assessment in our schools.

A particularly important influence in the changes we are experiencing is the growing recognition that teachers are often best placed to provide a true picture of the learning and achievements of their pupils. In the last decade or so the huge increase in teachers' interest in AfL has in turn prompted initiatives designed to explore how classroom assessments might be used for assessment *of* learning purposes. The comprehensive understanding of a pupil's learning progress can provide appropriate performance information to replace or work in conjunction with externally set and marked tests.

At this point a good assessment 'health warning' is needed. In some important quarters there is a worrying misappropriation of the notion of classroom assessment. A key element of AfL is for teachers to adjust their teaching to take account of what they learn from these assessments, including, for example, adjusting or improving their questioning and feedback. However, the other key issues are more explicitly pupil-focused. For example, experienced teachers continuously assess where pupils are in

their learning in order to help them to take charge of their own learning and to enable them to identify the next steps they should take. The teaching and learning focuses are mutually dependent, of course, but forms of didactic teaching, that is, with little or no engagement of the pupils, have not disappeared.

The health warning alarm therefore rings when the teacher – pupil interaction strays too far from a pupil-centred focus to a teacher-centred one that is a halfway house. It enables teachers to improve their teaching but may reduce the focus and explicit intention to improve learning *per se*. This is emphatically not a criticism of helping teachers to teach better but it does warn of the subtle dangers of changing focus from the pupil's learning to one in which the delivery of the teaching is the prime beneficiary. What is taught and how it is taught are still important elements in the service of better learning. The risk is that these become teacher-driven activities in which pupils play a largely passive role.

One manifestation of the potential problem would be the case in which pupils' knowledge and understanding are repeatedly assessed. For example, some 80 per cent of primary schools in England purchased 'progress' tests from the former National Assessment Agency in 2007 (BBC News, 2007b). Schools have a legitimate and worthy wish to monitor pupil progress as 'objectively' as they can, and a wide variety of commercial and state agency tests and assessment tasks is available for purchase. However, this is an unregulated and extremely varied practice that more often than not may constitute a significant overtesting of pupils across the UK.

There is also the strong potential for teacher assessment in the current Assessing Pupils' Progress arrangements for Key Stage 3 (QCA, 2009) to become merely a frequent check listing against criteria. The danger is that the pupil's control of the learning process may be lost to the gloss of beguiling phrases such as 'personalized learning' and 'diagnostic assessment'. The frequent, recorded assessments can be argued to be important feedback to individuals on their attainment, though the benefits of feedback on repeated failing might require a major stretch of the imagination. They can usefully diagnose weaknesses but if they remain a tool *only* for teachers to adjust their teaching, the process of learning may take a back seat.

Putting pupils at the centre of their own learning (clearly, no one else can do the learning for them) is a key element of AfL and the role of the teacher in this formative assessment process is a major focus of this book. So too is the role of the teacher in using the many classroom AfL interactions to draw up dependable reports (summative assessments) of a pupil's learning and achievements. The rise of such learning-oriented assessment is not restricted to the UK. It is a truly global phenomenon, charted for example in an international overview of formative assessment carried out

by the Organization for Economic and Co-operative Development (OECD, 2005).

## Formative and summative

At the outset of planning this book, we recognized that the terminology of assessment may present a challenge for many teachers and for some even the distinction between *formative* and *summative* assessment may generate a distracting degree of confusion. The simple but misleading distinction is that formative assessment is what a teacher might do during a period of learning and summative assessment is what they do at the end of it. This time-based distinction is unhelpful as it ignores the purposes and contexts of the assessments, and crucially how they relate to the learning process. We will develop the distinction later but it is important at this point to lay the ghost of 'time'-related assessment opportunities.

The most important argument is that an assessment activity is not inherently formative or summative; it is the purpose and use to which it is put that determines which of these it is. If a test is used to help pupils improve their learning, then technically it is being used formatively; it is assessment *for* learning. If an assessment activity and its 'result' (e.g. a score or a comment) is used in a report of a pupil's standard of performance, then it is being used summatively; it is assessment *of* learning. Logically then, an end-of-year test can be used summatively to report performance and formatively if it is also made the focus of pupil reflection and self-evaluation of their performance. But it is worth emphasizing that the term 'formative' only applies to the use of assessments to support learning.

When asked about assessment matters, many teachers might claim not to be 'assessment literate' though in truth they may be well-versed in good formative assessment practices such as effective classroom questioning and feedback. However, they are probably more accurate in their self-assessment when it comes to techniques that are used to make an assessment of learning through such approaches as class tests. While these will undoubtedly be useful in providing teachers with one type of measure of their pupils' grasp of a topic, their technical quality in terms of test design may be more miss than hit.

But existing competence and understanding of assessment matters is not necessary to benefit from this book. Our intentions are to inform and stimulate the understanding and development of assessment practices for individuals and schools, through reading the book and through professional development activities based on the assessment issues it covers. Capturing teachers' motivation and commitment to adopting any innovation invariably requires them to be convinced of its potential to

improve learning and raise standards. The process of adoption of new approaches also requires in-depth planning and commitment to ensure that any developments are undertaken in a high-quality and sustainable fashion. Subsequent chapters will therefore cover key professional learning dimensions such as the purposes of assessment, the need for evidence to support innovation, the process and steps to develop new practice, and perceptions of what counts as quality assessment in schools. Various perspectives will be considered as the book explains how teachers and schools should set about developing new practices, and how the system should react to support them.

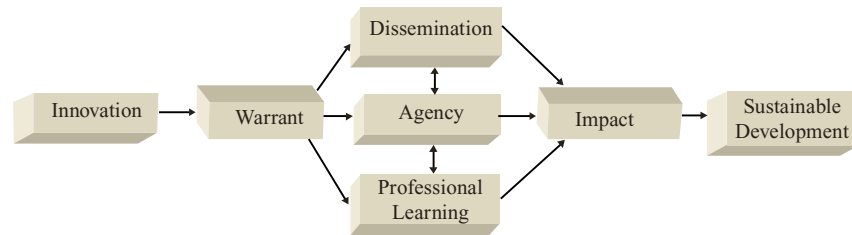
### **Key messages**

The chapters in this book draw partly on our extensive experience of international and national assessment development as members of the Assessment Reform Group and partly on our UK-wide review of most of the recent initiatives in teacher assessment in the UK. This work, the ARIA study,<sup>3</sup> was funded by the Nuffield Foundation and combined the views of 200-plus experts, from all four countries of the UK, who came together in a series of invited seminars to discuss the key issues emerging from the various initiatives. They included teachers and head teachers, academics, inspectors of schools, education consultants, school pupils, representatives of the four teaching councils and the professional teacher associations, representatives of the curriculum and assessment bodies, and both local authority officers and government representatives.

The most important feature of the book is therefore its breadth of perspectives on the finely focused issue of teacher assessment in schools. The book's main intention is to promote the development of teacher assessment by helping the main actors in schools, local authorities and government education and teaching agencies to capitalize on the insights from both the initiatives and expert participants in ARIA, and to better appreciate the processes that drive educational innovation and change.

Among the many insights that emerged from the ARIA study, there were two main conclusions. The first was that initiatives in assessment do not always take full account of the key dimensions of the change process or the needs of the communities involved. In this sense they can often be described as being underdesigned with consequent doubts, particularly in the sustainability of any new practices. It was clear that education systems, from school to national, must fully commit to all the necessary ingredients for sustained success if their objective is to promote and embed changes in assessment practice. The ARIA project presented these ingredients as seven key processes as set out in Figure 1.1. Note that this schema is not





**Figure 1.1** Key processes in ensuring sustainable change

intended to imply that the change process in schools and education systems is simple and linear. The individual elements interact in a complex, interwoven and interdependent manner, though arguably there has to be an initial innovation to begin the process and the ultimate aim is arguably to ensure that the aims are sustainable. These processes are considered in the chapters that follow.

The second conclusion reached in ARIA was that another phenomenon was acting across the scene. This might be summarized as many voices seemingly talking about the same issue (improvement in assessment practice) while using almost as many definitions of that issue as there are voices. Ultimately, such a situation dissolves into a *mêlée* of jargon used to describe different types of assessment, different uses of assessment and different perceptions of what is considered to be acceptable quality in assessment practice. Not having an explicit view of what is 'good' assessment for any particular purpose has the knock-on effect of making it difficult to decide what an improvement is. Indeed, it is also difficult to propose what effective dissemination or professional development for good assessment practice might look like.

A set of principles and standards is therefore needed to guide the development of effective assessment practice. These should be designed to enable any stakeholder group to assess the extent to which they are effectively promoting and sustaining desirable changes in assessment and its use. As with the model of key processes above, the principles and standards put forward in this book have emerged from the studies of recent projects and from the series of expert seminars. These are set out in detail in Chapter 2.

## Structure

The book is presented in three parts and Part I: *Lighting the fuse* is designed to introduce the context of exciting change that is empowering teachers

to add assessment skills to their repertoire of tools for improving pupils' learning. The chapters in this section therefore seek to ensure that readers are well prepared for dealing with what are sometimes quite complex concepts and practices. Following this introductory chapter, Chapter 1: *Assessment to support learning* covers the variety of purposes for assessment (including accountability), what teachers need to know about assessment and specifically the differences in purposes and usage that form the continuum between formative and summative assessment.

Following this, Chapter 2: *What is quality teacher assessment?* presents a synthesis of the insights and experiences of the 200-plus ARIA experts on a common language, in the form of principles and standards to guide assessment to improve learning in schools. Such an approach can also help to ensure that groups outside the school, such as advisory and inspection services, and policy makers, do not send contradictory messages on how assessment by teachers should develop.

The last chapter in the section, Chapter 3: *What's happening in the UK?* provides an overview of the recent curriculum-related assessment developments across the four main jurisdictions in the UK: England, Wales, Scotland and Northern Ireland. The analysis of the various policy developments confirms the widespread growth of teacher assessment and the many challenges still to be faced.

Part 2: *Spreading the word, widening the practice* begins the book's foray into understanding the complex processes involved in taking teacher assessment as a classroom innovation through to dynamically sustained practice. Chapter 4: *What is innovative about teacher assessment?* considers the concept of *innovation* in assessment and the importance of having evidence of teacher assessment being a viable and powerful classroom process (its 'warrant'). Chapter 5: *Moving beyond the classroom* considers the various models of *dissemination* available for promoting teacher assessment on a wider basis than an individual teacher's classroom. The focus of the chapter is on the pursuit of 'transformation' of practice rather than mere 'transmission' of good practices.

Perhaps one of the most important chapters then follows – Chapter 6: *Professional learning to promote teacher assessment*. The chapter emphasizes the importance of facilitating teachers with time for personal reflection and sharing experiences, and enabling them to participate fully in the design of effective *professional learning* activities. The problems of a superficial grasp of the key issues in teacher assessment, the misleading allure of classroom strategies and a failure to harness teachers' beliefs and commitment are also exposed. Chapter 7: *Teachers as self-agents of change* completes Part 2 by stressing the importance of *agency*, and particularly that teachers must be at the forefront of changing their own practices. As in Chapter 6, notions of top-down and bottom-up are used to argue

that teachers ultimately need more than information and guidance, or policy diktat, to undertake change. Their own needs and that of their pupils must generate their own intrinsic motivation to adopt and adapt new assessment practices in their classrooms.

Part 3: *Keeping it going* focuses on two important issues. The first, set out in Chapter 8: *Making a difference: evaluating the impact of innovations in assessment*, is that teachers and schools, and other interested parties, need to have a sound basis for believing that the changes are having an *impact*. Without this evidence of impact, there may be reluctance to maintain commitment and effort in the changes; indeed, there is every possibility that the teachers might query the point of doing it at all. Rather than simply perceiving there to be an impact, this chapter promotes a systematic approach to evaluating whether any changes are having the desired impact.

The final chapter, Chapter 9: *Embedding and sustaining developments in teacher assessment* is given over to the vital issue of ensuring that the changes in assessment practice are sustained. However, the key message here is that the needs of pupils, teachers and schools are constantly changing and that sustainability must therefore not be seen as a static situation. It is instead a dynamic concept, one that can only be secured if teachers are given the opportunity and encouragement to keep their assessment practices under review, as their needs and those of their pupils change over time.

## Using this book

Everyone involved in education, from teachers and schools through to policy makers, has to cooperate to ensure that desirable innovative practice is embraced and embedded throughout the system. Using this book, individual teachers and their schools as communities can develop, sustain and continually improve their assessment practices. In order to make the text immediately applicable in a school or local authority professional development setting, or in a higher education setting (e.g. Masters, PGCE, EdD), each end of chapter will offer a set of questions designed to prompt individual reflection and group discussion on the key issues concerned. In addition, a template for professional development workshops, which can address these types of question in the school, local authority or higher education context, is provided in Appendix 2. This proposes a 'snowball' format of individual reflection followed by discussion in pairs and four-somes, and provides suggestions on timing and approaches to engaging participants in discussions.

**Notes**

1. Throughout the book we use pupil, student and learner interchangeably, choosing the best fit with the context in each case.
2. The principles will be explained and developed in Chapter 2 – *What is quality teacher assessment?*
3. Analysis and Review of Innovations in Assessment – see Appendix 1 for more details.