Guidelines for Designing Programmes of Assessment

Introduction

The authors present these guidelines to be read with the following points in mind.

i) There is no linear order in the guidelines presented.

When reading the guidelines, you may not immediately come across those guidelines or important topics you would expect to be given priority. There is potentially more than one way of ordering the guidelines. As one example *costs* are important throughout the design process. However, because of the way this framework is constructed, *costs* are addressed near to the end.

ii) There is overlap between guidelines.

It appeared impractical and somewhat artificial to split every assessment activity into separate parts. The guidelines are highly related, and overlap and/or redundancy are almost inevitable. In the example of *costs*, which are primarily addressed as part of *cost-efficiency*, references to *costs* are actually made in several guidelines.

iii) The level of granularity is not equal for all guidelines.

Determining the right level of detail is a difficult endeavour, variable granularity reflects the fact that some issues seem more important than others, and others may have been investigated in depth.

iv) Assessment components and assessment information.

In the guidelines we have sought to find an overarching term that would cover all possible elements of the programme, such as assessments, tests, examinations, feedback, and dossiers. We wanted the guidelines to be broadly applicable, and so we have chosen the term assessment *components*. Similarly for outcomes of assessment components we have chosen assessment *information* (e.g. data about the assessees' competence or ability).

GENERAL GUIDELINES

I Decisions (and their consequences) should be proportionate to the quality of the information on which they are based.

II Every decision in the design process should be underpinned preferably supported by scientific evidence or evidence of best practice. If evidence is unavailable to support the choices made when designing the programme of assessment, the decisions should be identified as high priority for research.

III Specific expertise should be available (or sought) to perform the activities in the programme of assessment.

PURPOSE OF THE PROGRAMME

A1 One principal purpose of the assessment programme should be formulated.

A2 Long-term and short-term purposes should be formulated. But the number of purposes should be limited.

A3 An overarching structure which projects the domain onto the assessment programme should be constructed.

INFRASTRUCTURE

A4 Opportunities as well as restrictions for the assessment programme should be identified at an early stage and taken into account in the design process.

A5 Design decisions should be checked against consequences for the infrastructure. If necessary compromises should be made, either adjusting the purpose(s) of the assessment programme or adapting the infrastructure.

STAKEHOLDERS

A6 Stakeholders of the assessment programme should be identified and a rationale provided for including the expertise of different stakeholders (or not) and the specific role(s) which they should fulfil.

A7 The level at which various stakeholders participate in the design process should be based on the purpose of the programme as well as the needs of the stakeholders themselves.

PROGRAMME IN ACTION

Collecting Information

B1 When selecting an assessment component for the programme, the extent to which it contributes to the purpose(s) of the assessment programme should be the guiding principle.

B2 When selecting an assessment (component or combination), consideration of the content (stimulus) should take precedence over the response format.

B3 The assessment should sample the intended cognitive, behavioural or affective processes at the intended level.

B4 The information collected should be sufficiently informative (enough detail) to contribute to the purpose of the assessment programme.

B5 The assessment should be able to provide sufficient information to reach the desired level of certainty about the contingent action.

B6 The effect of the instruments on assessee behaviour should be taken into account.

B7 The relation between different assessment components should be taken into account

B8 The overt and covert costs of the assessment components should be taken into account and compared to alternatives.

B9 Assessment approaches that work well in a specific context (setting) should first be re-evaluated before use in another context (setting) before implementation.

B10 A programme of assessment should deal with error and bias in the collection of information. Error (random) is unpredictable and should be reduced by sampling (strategies). Bias (Systematic) should be analysed and its influence should be reduced by appropriate measures.

B11 Any performance categorisation system should be as simple as possible.

B12 When administering an assessment (component), the conditions (time, place, etc.) and the tasks (difficulty, complexity, authenticity, etc) should support the purpose of the specific assessment component.

B13 When scheduling assessment, the planning should support instruction and provide sufficient opportunity for learning.

PROGRAMME IN ACTION

Combining Information

B14 Combination of the information obtained by different assessment components should be justified based on meaningful entities either defined by purpose, content, or data patterns.

B15 The measurement level of the information should not be changed.

B16 The consequences of combining information obtained by different assessment components, for all stakeholders, should be checked.

Valuing Information

B17 The amount and quality of information on which a decision is based should be in proportion to the stakes.

B18 A rationale should be provided for the standard setting procedures.

Taking Action

B19 Consequences should be proportionally and conceptually related to the purpose of the assessment and justification for the consequences should be provided.

B20 The accessibility of information (feedback) to stakeholders involved should be defined.

B21 Information should be provided optimally in relation to the purpose of the assessment to the relevant stakeholders.

SUPPORTING THE PROGRAMME

Construction Support

C1 Appropriate central governance of the programme of assessment should be in place to align different assessment components and activities.

C2 Assessment development should be supported by quality review to optimise the current situation (Programme in Action), appropriate to the importance of the assessment.

C3 The current assessment (Programme in Action) should be routinely monitored on quality criteria.

C4 Support for constructing the assessment components requires domain expertise and assessment expertise.

C5 Support tasks should be well-defined and responsibilities should lie with the right persons.

Political and Legal Support

C6 The higher the stakes, the more robust the procedures should be.

C7 Procedures should be made transparent to all stakeholders.

C8 Acceptance of the programme should be widely sought.

C9 Protocols and procedures should be in place to support appeal and second opinion.

C10 A body of appeal should be in place

C11 Safety net procedures should be in place to protect both assessor and assessee.

C12 Protocols should be in place to check (the programme in action) on proportionality of actions taken and carefulness of assessment activities.

DOCUMENTING THE PROGRAMME

Rules and Regulations (R&R)

D1 Rules and regulations should be documented.

D2 Rules and regulations should support the purposes of the programme of assessment.

D3 The impact of rules and regulations should be checked against managerial, educational, and legal consequences.

D4 In drawing up rules and regulations one should be pragmatic and concise, to keep them manageable and avoid complexity.

D5 R&R should be based on routine practices and not on incidents or occasional problems.

D6 There should be an organisational body in place to uphold the rules and regulations and take decisions in unforeseen circumstances.

Learning Environment

D7 The environment or context in which the assessment programme has to function should be described.

D8 The relation between educational system and assessment programme should be specified.

DOCUMENTING THE PROGRAMME

Domain Mapping

D9 A domain map should be the optimal representation of the domain in the programme of assessment.

D10 A domain map should not be too detailed.

D11 Starting point for a domain map should be the domain or content and not the assessment component.

D12 A domain map should be a dynamic tool, and as a result should be revised periodically.

The term *blueprinting* is deliberately not used here, because this term is often used to denote a specific tool using a matrix format to map the domain (content) to the programme and the instruments to be used in the programme. With *Domain Mapping*, a more generalised approach is implied. Not only should content match with components, but the focus should be on the assessment programme as a whole in relation to the overarching structure (e.g. the educational curriculum) and the purpose.

D9 A domain map should be the optimal representation of the domain in the programme of assessment.
First of all, a domain map relates to the overarching structure (guideline A3). This domain map entails a more detailed specification of the overarching structure, including assessment components and content elements. A domain map is closely tied to the sampling of content and the strategies used (*Collecting Information*), and to combining information on specific content from different sources. This is related to the fact that a single instrument is never sufficient to claim that a particular domain is completely and validly assessed. A variety of assessment components is required, and these have to be mapped onto the domain. As such, *Domain Mapping* is part of the validation process and, in accordance to guideline B2, content prevails over format. Aspects to consider in describing the domain are the content (knowledge, skills, attitude), and the level of authenticity (simulated versus real). The programme should sample purposefully through content and levels of authenticity.

D10 A domain map should not be too detailed.
This guideline is formulated in order to avoid the pitfall of atomizing a programme of assessment into the smallest possible units of analysis. This would harm the integrative nature of a programmatic approach towards assessment. As such, a domain map should not contain too much detail, or too many dimensions (axes). Too much detail would diminish the degrees of freedom in assessment and frustrate the process of designing as well as administering the assessment programme.

D11 Starting point for a domain map should be the domain or content and not the assessment component.
In congruence with guideline B2 the assessment component (e.g. type or format) is a tool not a goal. The domain or content that should be measured is part of the goal. Often an assessment component is available or familiar to the user or designer and therefore becomes the first choice when designing an assessment programme. In some cases there will be a sound match between content and instrument, however, this is not guaranteed. Starting from the purpose and the nature of the domain to be assessed will focus the design process on achieving the purpose.

D12 A domain map should be a dynamic tool, and as a result should be revised periodically.
There is a risk that the domain map quickly becomes outdated, as virtually every field develops at a rapid pace. Also priorities and ideas change over time. Therefore the domain map should be updated periodically. The frequency with which updating is needed, depends on the context of the assessment programme and the pace of developments in the domain.

IMPROVING THE PROGRAMME

R&D

E1 A regular and recurrent process of evaluation and improvement should be in place, closing the feedback loop.

E2 If there is uncertainty about the evaluation, more information about the programme should be collected.

E3 In developing the programme (re-design) again improvements should be supported by scientific evidence or evidence of best practice.

Change Management

E4 Momentum for change has to be seized or has to be created by providing the necessary priority or external pressure.

E5 Underlying needs of stakeholders should be made explicit.

E6 Sufficient expertise about change management and about the local context should be sought.

E7 Faculty should be supported to cope with the change by providing adequate training

JUSTIFYING THE PROGRAMME

Effectiveness

Scientific Research

F1 Before the programme of assessment is designed, evidence should to be reviewed.

F2 New initiatives (developments) should be accompanied by evaluation, preferably scientific research.

External Review

F3 The programme of assessment should be reviewed periodically by a panel of experts.

F4 Benchmarking against similar assessment programmes (or institutes with similar purposes) should be conducted to judge the quality of the programme.

Efficiency: *cost-effectiveness*

F5 In order to be able to justify the resources used for the assessment programme, all costs (in terms of resources) should be made explicit.

F6 A cost-benefit analysis should be made regularly in light of the purposes of the programme. In the long term, a proactive approach to search for more resource-efficient alternatives should be adopted.

Acceptability: *political-legal justification*

F7 Open and transparent governance of the assessment programme should be in place and can be held accountable

F8 In order to establish a defensible programme of assessment there should be one vision (on assessment) communicated to external parties.

F9 The assessment programme should take into account superseding legal frameworks.

F10 Confidentiality and security of information should be guaranteed at an appropriate level.