Services for people with sight loss: how can we measure their effectiveness?

Thomas Pocklington Trust

Housing and support for people with sight loss

Report for Thomas Pocklington Trust

Authors: Graeme Douglas and Mike McLinden

Visual Impairment Centre for Teaching and Research (VICTAR)

Disability, Inclusion and Special Needs (DISN)

School of Education

University of Birmingham

Pocklington foreword

The report written by Graeme Douglas and Mike McLinden sets out some important information on outcome measures used when evaluating services for visually impaired people, including:

- An explanation of the often complicated language used when discussing outcome measures
- An overview of the key technical aspects of outcome measures, the relative benefits of different types and links to references which give further detail
- Things to consider when choosing an outcome measure
- Six case studies covering differing service types
- A list of outcome measures used in evaluating low vision services and a description of what they each set out to assess

The guide has been written for the following stakeholders:

- Professionals who want to evaluate the services they provide
- Researchers who are evaluating services and interventions
- Commissioners of research or services who are assessing the impact of their commissioned work

A forthcoming Pocklington Research Discussion Paper will provide a basic introduction to evaluation and a summary of this report.

Sarah Buchanan

Research Director

A guide to the design, selection and use of outcome measures in services for people with visual impairment

Report for Thomas Pocklington Trust

Date: Autumn 2012

Authors: Graeme Douglas and Mike McLinden

Visual Impairment Centre for Teaching and Research (VICTAR)
Disability, Inclusion and Special Needs (DISN)

School of Education
University of Birmingham

CONTENTS

INTRODUCTION

This Guide is one output from a project commissioned by Thomas Pocklington Trust. The aims of the project were to:

- explore the dimensions against which outcome measures for service providers in the field of visual impairment can be compared and assessed (e.g. relevance to population, validity, reliability, standardisation, etc);
- collate a summary of key outcomes measures currently used by a selection of services providers for people with visual impairment in the UK and assess these against the dimensions;
- produce guidance to help service providers and others determine which outcome measures may be most appropriate for measurement of different tasks/activities within the population of people with visual impairment.

The project was commissioned in a time of unprecedented interest in outcome measures for services for people with visual impairment in the United Kingdom. There is wider acceptance of the need to collect appropriate data in a climate of service-commissioning that seeks to identify, in a resource constrained environment, the difference that a service makes to people with sight loss. Even a cursory review of the literature shows a wide range of outcome measures currently in use within the broad visual impairment 'field'. There is a concern that many projects (research or evaluation, audit or review, children or adults) need to make use of such measures but, faced with a wide choice of possible tools they may be unsure which to choose, or how, or indeed whether to create their own. Rather than attempt to review this extensive literature base, or list the full range of outcome measures, we have drawn upon a number of relevant reports, articles and reviews and used these to illustrate key points. Whilst this approach is necessarily selective we hope it provides a helpful means of understanding the issues involved.

The Guide has been written to help 'stakeholders' in the field of visual impairment navigate the complex area of outcome measures by developing a helpful vocabulary of key terminology and concepts. Such stakeholders include professionals who are seeking to evaluate their practice, researchers who are assessing the relative value of interventions, and commissioners of research or services who are auditing the impact of commissioned work.

The Guide is divided into two parts:

Part 1 – Definition and selection of outcome measures

- Definition of outcomes and outcome measures;
- Outcome measures – some technical considerations;
- Criteria for the selection of appropriate outcome measures.

Part 2 – Example outcome measures

- Illustrative case studies;
- Collation of some key outcome measures.

The Guide does not develop discussion about the range of evaluation, audit and review methodologies that services and research may use (and in which different outcome measures are applied). A wide range of generic materials exist to guide readers on that topic and Pocklington has produced a brief overview to signpost sources of particular interest in the evaluation or measurement of service outcomes.

We hope that this Guide goes some way to ensuring that service providers and other stakeholders are better informed about the options available to them in relation to the purpose, design and use of outcome measures and as such are able to draw on a more robust evidence base as they report the impact of their work.

We would like to take this opportunity to thank Thomas Pocklington Trust for commissioning the project and in particular Sarah Buchanan for her invaluable input. We hope you find the Guide of value and we welcome feedback to assist in revising a future edition.

Graeme Douglas and Mike McLinden,
VICTAR, 2012

PART 1 – DEFINITIONS AND SELECTION OF OUTCOME MEASURES

1 Definition of outcomes and outcome measures

This Guide aims to develop a helpful vocabulary which enables stakeholders to navigate the complex area of 'outcome measures'. Key points we introduce in this section:

- In relation to service provision, a useful distinction can be made between 'outcomes' and 'outputs'.
- It is helpful to think of 'outcomes' as referring to the changes that take place as a result of a given service (intervention, activity, programme or input).
- 'Outcomes' can refer to changes that take place at different levels (e.g. an individual, a group or a service level).
- 'Outputs' are associated with particular 'products' or services the provider offers (e.g. number of people using the services in a given year rather than any change associated with use of the service). These outputs can also be described as 'activities'.
- Whilst some outcomes may be clearly linked to particular outputs, not all will be.
- Outcome 'measures' are the tools that can identify and measure any changes that result from an intervention.
- An important distinction is made in this Guide between primary and ultimate outcomes, and in turn primary and ultimate (sometimes referred to as 'secondary') outcomes measures.
- Evaluation of services involves establishing change which has resulted from that service (i.e. outcomes). A key part of such an evaluation is to identify the primary outcome(s) of the service and establish a measure which is appropriate for that outcome.
- Appropriate 'alignment' of the primary outcome and outcome measure is a important feature of an effective evaluation.

1.1 Outcomes

There is broad agreement in the literature that outcomes refer to 'changes' that take place as a result of a particular activity, programme or input. As an example, Burns and Cuppitt (2003) describe outcomes as 'the changes, benefits, learning or other effects' that happen as a result of particular activities by an organisation, for example, improved confidence or increased skills (p4). This is supported by Myers and Barnes (2005), in relation to children who describe outcomes as the changes that have been made as a result of a given programme's activities. They note that outcomes have the "power to answer the question 'What difference is one particular service making?'" (p3), arguing that outcomes are important because they provide a mechanism by which services (or programmes) can assess the effect that they have had on their beneficiaries:

> After describing the implementation and process of delivering services, at some point programmes and services need to produce evidence to document what they have realized for the populations with whom they have been working. That way, observers of the programme are able to attribute value to the work that has been undertaken. (p 5)

'Outcomes' and 'outputs' differ. As an example Burns and Cuppit (2003) refer to outputs as the detailed activities, services and products of an organisation (e.g. key work sessions, group work sessions, or advice and information). It is useful then to think of outputs as relating to the products or services an organisation offers or provides to it users, whilst outcomes refer to the difference a given service has had on the service users. As an example, we can consider a weekly class set up to help elderly people with a visual impairment develop a more active lifestyle. An example of an output measure would be the number of people regularly attending the class (measured for example through the use of a weekly register), whereas an example of an outcome would be changes the class has made to the service users activity levels. An outcome measure would then be a method of measuring this change in activity, e.g. a questionnaire, interviews or, depending on the focus, objective performance measurements of their fitness levels.

Burns and Cuppitt (2003) distinguish between "outcomes for individuals" and "outcomes for communities" (i.e. those drawn upon for policy change). Although the outcomes are related, they note that outcomes can occur at many levels including:

- individual clients
- families or households
- the community
- the environment
- organisations
- policy

A similar distinction is made by Myers and Barnes (2005) in describing outcomes that can be:

- changes in the people with whom the programme comes into contact
- changes in the organisation with which the programme comes into contact
- changes in the environment in which the programme operates

Colenbrander (2010) makes an important distinction between primary outcomes and ultimate outcomes. The ultimate outcome of a service is linked to its broader aims and purpose including for example, improving people's quality of life. In contrast, the primary outcome of a service is the more specific change the service is seeking to bring about. The distinction is important because it makes a difference not only to how outcomes are conceptualised, but also to how they are measured.

1.2 Outcome measures

If outcomes are viewed as referring to 'change' as a result of a given intervention, outcome measures can be linked to the evaluation of this intervention in order to 'detect' any change. Given the range of outcomes that can be measured, a range of approaches are outlined in the literature that can be used to measure or 'detect' changes in relation to individuals with visual impairment.

Robson (2002, p.202) describes an evaluation as:

> An attempt to assess the worth or value of some innovation, intervention, service or approach.

This is a useful statement to draw upon, as it lays out key parameters for an approach to an evaluation. Evaluations of interventions can, of course, seek to understand many aspects of how that intervention worked (e.g. the amount of time it took, how much it cost, who was involved, and client satisfaction/feedback). Traditionally services have also been interested in 'standards' , which means that the services provide a quality of service which is acceptable (e.g. procedures are followed, services are delivered in a timely manner to the intended clients, the desired services are delivered). Standards can provide criteria against which evaluations can take place and can define the primary outcomes sought. National policy documents exist which define standards for service providers, e.g. DfES (2002) present guidance quality standards for use by education support services for children and young people with visual impairment; ADSS (2002) present national standards of social care for visually impaired adults. Whether standards are guidance or mandatory, they offer service managers, commissioners and clients guidelines as to what might be expected from the services in question (whether rehabilitation, visiting teacher, or any other service of interest).

To measure any change following an intervention, the assessment approach used should be aligned to the outcomes that need to be identified. This is captured succinctly by Myers and Barnes (2005) in noting that once "outcomes have been identified it makes the evaluator's task easier by being able to match the approach and method to more reliably measure the anticipated changes" (p18).

Much of our discussion about outcome measures in this Guide is limited to a particular approach which is commonly reported in the literature – the use of questionnaires (either self completion or completed through an interview). However, it is important to note that outcome measures do not have to be limited to these methods. Other, more direct and objective methods can be appropriate, such as observation of mobility performance or timed performance of activities of daily living (e.g. Colenbrander, 2010). Education, in particular, has a tradition of using performance assessments as outcome measures (e.g. reading performance, test/examination performance).

1.3 Primary and ultimate outcome measures

As well as a distinction between outputs and outcomes, it is also useful to think about the difference between primary and ultimate outcome measures. Colenbrander (2010) notes that primary outcome measures are directly linked and relevant to the service / intervention being evaluated. Ultimate (or 'secondary') outcome measures tend to be more global and relate to improving people's 'quality of life'. An important task when designing a review, audit or an evaluation of an intervention is to ensure that appropriate primary outcomes and measures are identified (and to ensure that ultimate outcomes and measures are not unwittingly used). We can return to the adult day class example to consider this distinction further.

If we assume that the primary outcome of the class was to 'improve the activity levels of the service users' then this should form the basis for deciding what is the most appropriate outcome measure/s to use (i.e. to determine whether there has been any change in their levels of activity). The ultimate outcomes of this class relate to broader improvements in the service users reported quality of life as a result of attending the classes over a period of time. As such, as well as a change in reported activity levels, there may well be other changes that service users report as result of attending the day class (e.g. feeling more confident about meeting new people, feeling more able to participate in other physical activities) – these ultimate outcomes will however require different outcome measures to capture any reported change.

The focus of Colenbrander's review was vision rehabilitation / low vision services work and he identified 'functional vision' as the primary outcome in that discipline. Because vision rehabilitation has broad application, functional vision primary outcomes can be usefully divided into more precise measures (e.g. visual reading, visual motor skills), many of which are used as sub-scales of 'vision-related quality of life' outcome measures.

Colenbrander's language can be used for other services, e.g. befriending services might have a primary outcome (and measure) linked to 'social contact'; mobility training services might have a primary outcome (and measure) linked to 'independent journeys' taken. Services which have a broader remit such as employment services may well have a primary goal for outcomes related to 'employment activity' (e.g. getting and maintaining a job, making job applications), but are also likely to have other primary outcomes

and measures linked to 'preparation for employment' (e.g. mobility, ICT skills). Examples of how these primary outcomes might translate into particular outcome measures are presented in Table 1.

Table 1.
The relationship between services, primary outcomes, primary outcome measures, and ultimate outcomes

Nature of service provided	Example of primary outcome(s) of service	Example of primary outcome measure	Example of 'ultimate' outcome
Vision rehabilitation/ low vision work	To improve a client's use of functional vision	Measures of functional vision before and after programme. e.g. measured by 'Low Vision Quality-of-life Questionnaire (LVQOL)'	Improved quality of life generally, improved mobility, improved access to information.
Befriending service	To widen the scope of a client's social contacts	Measures of the numbers and quality of social contact experienced by client, e.g. a questionnaire exploring social contact each week.	Improved sense of inclusion, drop in feelings of depression
Mobility training service	To increase the number of independent journeys made, and improved confidence in mobility	Measure of the numbers of independent journeys undertaken each week from home, e.g. a log of journeys made. Increased confidence in mobility as measured by 'Patient-Based Assessment of Difficulty in Mobility'	Increased social contact, reduced sense of dependence, greater readiness for employment
Employment services	To gain and maintain employment	Changes in employment activity: applications for jobs, interviews, voluntary work, employment	Improved quality of life generally, greater economic security

1.4 Aligning outcomes to service provision

Table 1 highlights the importance of ensuring alignment between primary outcomes and outcome measures and suggests a way of thinking about how to categorise the range of measures according to the nature of the service provided.

A general definition of 'outcome measures' can seem rather bland and unhelpful because it does not make explicit links to the topic of interest. More specific definitions can be achieved by considering two inter-related areas. First, consideration of outcome measures which have particular relevance to visual impairment (e.g. mobility,

access to information and literacy, access technology, and use of low vision devices). Second, the nature of the service provided.

For the purpose of this Guide, a broad distinction is made between 'education', and 'rehabilitation, social care and low vision training'.

1.4.1 Education

Education often draws upon summative assessments of attainment in particular areas of the curriculum as one form of outcome measure. In a review of international literature in relation to visual impairment education, Douglas et al (2009) made a key distinction between the 'curriculum' and the 'additional curriculum'. Unsurprisingly, the education sector has clear outcome measures associated with the curriculum – most obviously educational assessments such as national examinations (e.g. GCSEs and A-levels). With some modification young people with vision impairments are able to be included in these assessments, and it is possible to use these outcome measures in considering educational outcomes for visually impaired pupils (e.g. Chanfreau and Cebulla, 2009). The review highlighted other areas of the curriculum which are of particular concern, most notably print and braille reading. Again, those concerned with education draw upon adapted assessments (e.g. modified tests of reading ability) to measure children's progress in these areas.

The additional curriculum can be thought of as either 'over and above' the mainstream curriculum, or consisting of areas which are outside the mainstream teacher's expertise and require the input from professionals with specialist training/knowledge (e.g. specialist teachers, mobility officer). The areas identified in the Douglas et al (2009) review were: 'mobility and independence', 'social and emotional inclusion', 'access technology', and 'low vision training'. In comparison with the mainstream curriculum, the review suggests there is less of a tradition of drawing on summative assessments of attainment in these additional curriculum areas. Nevertheless, educators do and should try to use outcome measures for these important outcomes.

Another possible area of educational outcome which was not covered by the Douglas et al (2009) review is that of 'destination' beyond school. Typically, these types of outcome measure are related to the labour market (e.g. in employment, in education or training, or none of these - so-called NEET).

1.4.2 Rehabilitation, social care and low vision training

'Rehabilitation, social care and low vision training' is a wider and less well-defined area than education. Perhaps this is unsurprising given this area covers a range of ages and circumstances. Nevertheless, broad conceptions of specialist areas of concern in the area of visual impairment have been identified (e.g. in the UK, ADSS 2002 and Low Vision Working Group 2007). These conceptions are reflected in the services which have been developed over many years by statutory and voluntary organisations and also described by key commentators, e.g. the American Foundation for the Blind (AFB) in the USA. AFB has published a series of influential books (the "Foundations" series) which describe aspects or rehabilitation and social care. The disciplines include:

- orientation and mobility (e.g. Weiner, Welsh and Blasch, 2010),
- counselling and employment (e.g. Graves, Moore and Patterson, 1997)
- communication (including literacy and computer technology) (e.g. Ponchillia and Ponchillia, 1996)
- activities of daily living (including food preparation, personal care) (e.g. Ponchillia and Ponchillia, 1996)
- low vision training / vision rehabilitation (e.g. Corn and Erin, 2010)

1.4.3 Links with contemporary strategies – the UK Vision Strategy

The UK Vision Strategy (2008) was developed by a wide-ranging alliance of statutory health and social care bodies, voluntary sector organisations, eye health professionals and service users with an aim of developing services for visually impaired people in the UK. Related documents have been developed in relation to implementing the strategy (e.g. UK Vision Strategy Implementation Plan for England, 2009; Seeing It My Way, 2011). 'Seeing It My Way' (2011) is described as a universal quality and outcomes framework for blind and partially sighted people. The focus on 'outcomes' rather than service delivery is of particular relevance to this Guide:

> "Seeing It My Way is presented as outcomes rather than services because delivery of these outcomes and the organisations that provide them will vary across England, Scotland, Wales and Northern Ireland" (p5).

The ten 'outcomes' in the framework were updated in 2012 following consultation (UK Vision Strategy, 2012) and are presented in Table 2, with cross-reference made to traditional areas of service described above. In the context of this Guide these outcomes can be considered to be aspirational rather than true 'primary outcomes' and, as such, provide a framework for 'ultimate outcomes' which services should be seeking to address and contribute towards.

Table 2.
Outcomes from 'Seeing It My Way' (2012)

'Seeing It My Way' outcome	Service area
1. That I understand my eye condition and the registration process	Information support services; Eye clinic liaison
2. That I have someone to talk to	Communication; Social isolation
3. That I can look after myself, my health, my home and my family	Activities of daily living
4. That I receive statutory benefits and information and support that I need	Information support services
5. That I can make the best use of the sight I have	Low vision training
6. That I can access information making the most of the advantages that technology brings	Communication
7. That I can get out and about	Mobility
8. That I have the tools, skills and confidence to communicate	Communication
9. That I have equal access to education and lifelong learning	Education
10. That I can work and volunteer	Employment

Of particular note is that these outcomes are described from the perspective of the 'service user' to emphasise the important philosophy of 'nothing about me without me' (Seeing It My Way, 2011, p6), with each outcome then expanded through a 'definition'. An example of the definition provided for Outcome 2 is presented in Table 3.

Table 3.
Example of how an outcome from 'Seeing It My Way' is defined

Outcome 2 – That I have someone to talk to

- I want help to come to terms with my condition and to help me address the impact of sight loss on my life. I want that support to be appropriate to my needs, whether it is advice or extended professional counselling. I want to be given dedicated time and support to help me achieve personal well being. This means that:
- I will be offered advice and appropriate support at the point of diagnosis by a trained professional or information service.
- I will be informed about available services to support my emotional needs and if necessary, be helped to access them at the right time and at the right place with acknowledgement that as my condition changes my need for such support might be ongoing.
- I will have access to a well facilitated peer support group.
- If I have additional disabilities, please remember that I may not be able to tell you how my sight loss makes me feel. I may need an advocate.

The links to service areas discussed above (e.g. Table 1) reminds us that while the conceptualisation of 'outcomes' is an important contemporary take on the evaluation of services, the underlying service areas are nevertheless familiar.

1.4.4 Links with outcomes defined by funders

An important factor which may dictate how outcomes are conceptualised is that some funders of services may prescribe the outcome measure that should be used in service evaluation. In some cases funding may even be conditional upon results measured against this outcome measure (a 'payment by results' approach to funding). A particular example of this is in relation to some current government funding of employment services where payment is only made when clients have entered employment for a sustained period of time. Similarly, some funders emphasise measurable outcomes and funding applications need to be mindful of this approach (see for example Big Lottery Fund, 2004 – 'Your project and its outcomes'). Such an emphasis will help determine the choice of outcome measure and in some cases it can dictate the design of a given service (and define the primary outcome of that service).

1.5 'General' and 'specific' outcome measures

We noted above that outcome measures can usefully be considered as aligned to specific areas of targeted change in visually impaired people's lives, which in turn are linked to targeted interventions (e.g. mobility, low vision aids, counselling). In spite of the importance of discipline-specific outcome measures, some general approaches to outcome measures are relevant to this discussion (and have been commonly used). Perhaps the most relevant example here is measuring perceived 'quality of life' (e.g. through the use of well-being / self-esteem measures).

Arguably, considering 'quality of life' outcome measures offers a more holistic approach. After all, services will usually have a broad aim (or 'ultimate outcome') of improving people's lives. It seems reasonable therefore that service evaluation should seek to measure such improvements. A number of attempts have been made to develop quality of life measures (e.g. the questionnaire-based 'self-perceived quality-of-life scale'), and their application has been much broader than evaluation of services (e.g. comparison of wellbeing across countries). Given that quality of life measures seek to assess general well-being however, it might be regarded as 'ambitious' that a service should hope to make a demonstrable difference to clients' overall quality of life. Indeed, in a review of outcome approaches to vision rehabilitation, Colenbrander (2010) notes that "asking global, generic questions about quality of life is not sensitive enough" (p165). This suggests that generic quality of life measures should be used cautiously as outcome measures as they may not align with the primary outcomes of the service being evaluated.

An example of the limitations of drawing exclusively on a generic instrument in relation to children is reported in a study by Van Dijk et al (2007). The study used self-report and proxy report versions of KIDSCREEN-53 to measure the health-related quality of life (HRQoL) of young people who had survived retinoblastoma in the Netherlands. The limitations of drawing on a general HRQoL questionnaire are reported with the authors noting that a general instrument such as KIDSCREEN only measures:

> broad areas of HRQoL and may not identify such issues specifically associated with retinoblastoma. In future studies it is advisable to use a retinoblastoma-specific instrument or a vision-related QoL instruments besides a general HRQoL instruction. Unfortunately there is a shortage of such instruments, which in itself presents a challenge for future research (p7).

Partly related to this challenge, dedicated 'vision-related quality of life' measures have been created to include items which are more specific to some of the service areas of particular interest in the field of visual impairment identified above (e.g. mobility, accessing information, activities of daily living). Some of these are listed in Part 2 of the Guide.

2 Outcome measures – some technical considerations

In this section we are concerned with the technical aspects of outcome measures and examine a number of the commonly used terms which are used to describe the properties of a particular measure. For the purpose of this discussion, it is helpful to make reference to the term 'psychometrics' which is used to describe the construction and development of psychological measurement (for example, questionnaires, personality tests etc). We examine first the psychometric properties of validity and reliability and how they relate to the notion of an 'ideal' outcome measure. We then consider these 'ideal' properties in the practical context of selecting an outcome measure which can be considered to be 'good enough'. Linked to these more practical and contextual concerns we briefly examine the ways in which outcome measures can be used and introduce the notion of 'evaluation design'.

A challenge of writing a Guide such as this is that on one hand it requires the necessary technical detail to unpick some complex issues, while on the other hand the technical detail can be intimidating for readers who are less familiar with research. The following brief overview of the 'psychometric properties of a good outcome measure' is a case in point. For the interested reader this overview provides an introduction to some key concepts, and offers links to references where more detailed information can be found. Alternatively, other readers may prefer to only read the key points below before moving on to the section 'Relative qualities of different outcome measures'.

Key points introduced in this section:

- The key psychometric properties of validity and reliability.
- The notion that a valid measure can be considered to be one which measures what it purports to measure. Approaches and techniques introduced are: content (or face) validity; construct validity; and Rasch analysis.
- The notion that a reliable measure can be considered to be one which is stable and consistent. Approaches and techniques introduced are: internal consistency and test retest reliability.

- As few outcome measures are standardised (i.e. have all the 'ideal' psychometric properties), other pragmatic factors must also be considered when selecting an outcome measure.
- Appropriate 'alignment' of the primary outcome and outcome measure, i.e. to ensure the outcome measure is directly relevant to the service being evaluated.
- Ideally an outcome measure which has good reliability and validity would be selected, i.e. it has been 'standardised' in some way and is reported in research literature.
- If this is not possible, an outcome measure should be considered which follows a 'standard' format.
- It can be appropriate to create bespoke outcome measures based upon satisfaction or rating questions which are specific to the service being evaluated.
- The evaluation design is the procedure followed when collecting data in a rigorous way. Examples of different evaluation designs are introduced.

2.1 The psychometric properties of a good outcome measure

As noted above, the term psychometrics is concerned with psychological measurement. Using the theories of psychometrics it is possible to analyse the properties of an outcome measure and make judgements as to whether it can be described as being a 'good' measure. For example, Massof and Rubin (2001) reviewed the validity and reliability of 13 functional vision questionnaires (or outcome measures) in this way. They note that many of the questionnaires followed a similar developmental process which included:

- Interviewing a sample of participants to identify their functional complaints.
- Developing an initial set of items (a broad item pool).
- Consulting experts in the field to refine and 'validate' the selection of items.
- Developing a response scale for each item.
- Administering pilot instrument to a sample of participants.
- Reviewing and refine the instrument.

Tests which have undergone rigorous development and have psychometric properties are often said to be 'standardised'. In this section we are mainly concerned with examining terminology relating to the psychometric properties of an instrument, with a particular focus on validity and reliability.

2.1.1 Validity

A 'valid' measure can be considered to be one which measures what it purports to measure, that is, the measure generates a value which is considered to be trustworthy.

An important term to consider when examining the properties of a given outcome measure is that of 'validity'. This is a large topic area which incorporates many other aspects and as such can be overwhelming. However, at a broad level, the validity of an instrument can be described as 'the extent to which an instrument measures what it purports to measure' (Margolis et al 2002, p 796). Similarly, Massof and Rubin (2001) refer to validity as being 'a statement of confidence that the instrument accurately measures what it intends to measure' (p533).

It can be useful to think about validity in relation to trustworthiness. In short, will the information you have gathered (or plan to gather) provide information that will convince others? (i.e. to what extent can this information be considered to be trustworthy?). This point is captured succinctly by Massof and Rubin (2001) in noting that validity is 'not a statement about the measurement operations, but a statement about the interpretation of the instrument's score.' (p 533).

We draw on one commonly used instrument in clinical practice, the Vision Function-14 (VF-14), to examine the term further. The VF-14 is described as an index of functional impairment in patients with cataract (Steinberg et al 1994). The 14 items are addressed by 18 questions that relate to vision dependent activities performed in everyday life that can be affected by cataract (e.g. reading text, traffic signs, playing games, cooking, driving). The instrument was developed and evaluated in a group of 766 patients undergoing cataract surgery for the first time (i.e. it was 'standardised'), and is described as 'a reliable and valid measure of functional impairment caused by cataract and provides information not conveyed by visual acuity or a general measure of health status' (Steinberg et al 1994).

Based on the description above, having a valid measure means that the instrument has been shown to measure what it purports to measure. In this case there is evidence to show that it measures the functional limitations in vision-dependent activities performed in everyday life that can be affected by cataract. In essence this means that the instrument generates a value that can be considered to be trustworthy – but crucially, only in relation to individuals who have cataract (as these were the participants the instrument was 'validated' with). If we wanted to ensure the instrument generates a value that can be considered to be trustworthy when looking at the limitations of vision dependent activities affected by other conditions we would have to 'validate' the instrument with a different group of participants, e.g. Linder at al (1999) did this in relation to people with retinal disease. For these reasons Margolis et al (2002) reported the VF-14 was one of the few instruments that had 'undergone extensive psychometric testing and in general appear to perform fairly well with varying ocular conditions' (p808).

With this broad understanding of validity we consider briefly how a researcher actually sets out to ensure that a measure is considered to be 'valid'. To more fully understand the process we need to consider briefly a number of additional terms relating to validity. Our particular focus is on 'content' and 'construct' validity, as well as a term that is gaining increasing recognition as being important in instrument development, namely Rasch analysis.

Content validity (or face validity)

At a basic level an outcome measure may have 'face validity'. This refers to users of the measure (or other 'experts' in the field) agreeing that the measure seems reasonable. In essence content-related evidence is used to justify the selection of items in a given instrument and the response scales (Massof and Rubin 2001). In relation to vision specific measures for assessing health-related quality of life, Margolis et al (2002) note that content validity is 'the degree to which the instrument is reflective of aspects important to the patients and disease of interest. In other words, does the instrument reflect patient concerns and interests?' (p 796). In practice as reported by Margolis et al (2002) whilst designers of questionnaires may not report on content validity per se, they may describe the methods that were used during their development to ensure face validity with the participants in the study and the experts.

Taking the VF-14 as an example, Massof and Rubin (2001) note that 'considerable effort' (including interviewing patients and ophthalmologists and through a literature review) went into choosing the 14 items to be included in the instrument to ensure that it reflected areas that were important to people with cataracts.

Construct Validity

Construct (or construct-related) validity is a difficult term to describe but in essence it refers to how well an instrument correlates with other indicators of similar or related constructs (Margolis et al 2002). Put very simply, one can ask 'does the outcome measure agree with other measures which are trying to evaluate similar things?' If it does, then it is said to have construct validity. This can be circular because different measures are used to validate one another – in fact, it can be very difficult because alternative measures might not exist (that is why researchers often only claim 'face' validity). In their review of 22 vision-specific instruments assessing visual functioning and/or impact of visual impairment on quality of life or daily activities, Margolis et al (2002) report that for the instruments reviewed, 'construct validity was assessed using correlations with objective clinical indicators of visual function, other vision-specific instruments, or generic HR-QOL instruments' (p796).

Rasch Analysis

Many outcome measures ask respondents to score or rate their agreement with a statement or amount of difficulty with a task. Adding these scores creates an overall score for the outcome measure for that participant either as a single score or as subscales. Limitations of this approach mean that the overall scores are problematic because they do not fit on an interval scale. An interval scale is one in which the scale has equal increments, i.e. "spaced evenly along a difficulty axis" (Colenbrander, 2010). For example, if an interval scale ranges from 1 to 5, then the difference between 1 and 2 is considered to be the same as between 2 and 3, 3 and 4, and 4 and 5. If the outcome measure does not have an interval scale (for example, a scale that does not have equal increments), combining participants' scores to create averages across participants should be done with caution e.g. Massof and Rubin (2001).

Rasch analysis is a statistical technique which seeks to overcome this problem. The analysis serves to calibrate an outcome measure

scale so that it is has an interval scale. Whilst the inclusion of a Rasch analysis in the process of instrument design is often seen as a 'gold standard' in outcome measures (e.g. Colenbrander, 2010; Massof and Rubin, 2001; de Boer et al, 2004), many do not incorporate this approach. For example, Massof and Rubin (2001) reported that a Rasch analysis had been incorporated into the development of only three of 13 reviewed measures. In actuality, many test designers and researchers either choose to use existing outcome measures on pragmatic grounds (e.g. test availability or cost of test development) or ignore the issue.

2.1.2 Reliability

A 'reliable measure' can be considered to be one which is stable and consistent. A 'reliable outcome measure' would be expected to provide the same value when carried out with the same person in the same circumstances.

Reliability describes the consistency of an outcome measure's score when re-administered (test-retest reliability) or the consistency of the scores across different conditions of administration (e.g. effects of environment, person administering the instrument, or mode of administration) (Massof and Rubin 2001, p534).

A number of tests are used to provide evidence of the measurement error (or reliability) associated with an outcome measure. As noted by Massof and Rubin (2001), most reliability tests are designed to 'assess the precision of the instrument' (p533). We consider below two commonly used measures of reliability (Margolis et al 2002): 'internal consistency' and 'reproducibility' (or 'test-retest' reliability).

- **Internal consistency** reliability is the extent to which the items in a given outcome measure assess the same construct (Margolis et al 2002) – that is how consistent the items are internally. This consistency is often determined through the use of a statistical test called 'Cronbach's formula' which gives a 'Cronbach's Alpha' value. A value for a given outcome measure of 0.7 or greater is commonly reported to indicate that a test has acceptable internal consistency.
- **Test-retest reliability** is probably the measure most people associate with the term reliability. It 'refers to the degree to which the scores remain the same over time when no change is expected' (Margolis et al, 2002, p 796). Tests of this type of

reliability include types of correlations (e.g. Kendall's index of concordance, the Kappa coefficient). However, Massof and Rubin (2001) noted that this was rarely reported for the development of the 13 instruments they reviewed.

2.2 Relative qualities of different outcome measure

The psychometric properties outlined above are clearly important when selecting an outcome measure to use in a service evaluation (and particularly for deciding if a measure can be considered to be 'trustworthy' and/or 'reliable'). Indeed, consideration of such psychometric properties provides a means of reviewing the relative merits of different outcome measures in the literature (e.g. Margolis et al, 2002; Massof and Rubin, 2001). However, other considerations are also important when deciding if an outcome measure is 'good enough' for a given purpose. As noted above, few outcome measures have all the psychometric qualities which might be desired. There is a need therefore to be pragmatic. Central to this is the notion of 'alignment' (i.e. linkage between the primary outcome and outcome measure) to ensure the outcome measure is directly relevant to the service being evaluated. Margolis et al (2002) sum up the importance of this alignment succinctly in noting: "It is important to explicitly state the desired outcomes and ensure that it is consistent with the selected [outcome measure]" (p808-809).

A key source of standardised outcome measures is research literature in relation to the service area of interest. As an example, a rich source of information is a review by Bins et al (2009) that provides an extensive list of outcome measures (and their associated psychometric properties) in relation to low vision services. Drawing on this review, we have provided a list of outcome measure in Part 2 of this Guide that is cross tabulated with what we judge to be associated primary outcome areas. Similarly, in the illustrative case studies in Part 2, three examples are given where researchers have chosen standardised outcome measures which have been designed by others (i.e. Case Study 2 in relation to eye clinic support services, Case Study 4 in relation to low vision services, and Case Study 5 in relation to counselling services). In each of these case studies, the researchers chose outcome measures with 'good' psychometric properties in order to strengthen their evaluations.

If a relevant standardised outcome measure does not exist however, then it is possible to develop one. As an example, in illustrative Case Study 1 the researchers were dissatisfied with available outcome

measures in relation to low vision rehabilitation and children and chose to develop their own. However, this process of development is relatively time consuming and resource intensive.

Other outcome measures may have been developed which are relevant and can often be identified in research and evaluation literature. These measures often have good 'face validity', i.e. they appear to be meaningful and trustworthy because the questions seem relevant and others with similar interests believe them to be useful. In the illustrative Case Study 6 (an intervention in relation to literacy / reading performance) the researchers made use of an existing outcome measure (the 'Reading Behavior Inventory', or RBI). In fact, the repeated use of an outcome measure in order to collect data is an important part of standardisation (therefore in the case of the RBI example, its repeated use in evaluations will lead to improving its psychometric properties). Related to this, in Case Study 4 the researchers were dissatisfied with available outcome measures in relation to measuring the impact of specialist visual impairment employment services and therefore designed a new pilot assessment tool. Their aim was to construct a standard tool for use within their organisation to measure distance from the labour market. As part of a future phase of the study, it is anticipated that enough data will have been collected to enable certain psychometric qualities of the assessment tool to be developed.

These examples are drawn upon to illustrate that it is a rather 'grey area' in deciding when an outcome measure is 'standardised' or 'not standardised'. Indeed, it is probably more useful to consider the psychometric properties of an outcome measure (e.g. to determine whether it is valid and/or reliable) and that it follows a 'standard' format. A standard format is important because it ensures the outcome measure is consistently used irrespective of the measure administrator (or the time and place of administration). Such consistency is an important part of any meaningful outcome measure.

A further alternative is the creation of bespoke outcome measures based upon satisfaction or rating questions which are specific to the service being evaluated. Adopting a standard format remains important to ensure the outcome measure is used consistently. An example in relation to low vision rehabilitation is presented by Aspinall et al (1999). The researchers constructed satisfaction questions which were linked to particular aspects of a hospital-based low vision service under evaluation. The remit of the research was broader than the satisfaction patients had with the service,

rather it investigated how reported satisfaction correlated with a number of other patient characteristics irrespective of the service they had received (e.g. general quality of life, wellbeing, religiosity). Nevertheless, the construction of the service satisfaction questions is an interesting example of a practical and small scale approach to evaluation (particularly as the authors reflect upon the pros and cons of using satisfaction scales for service evaluation purposes).

Finally, in addition to the selection of the outcome measure itself, and the notion of using a standard format to ensure consistency, is the 'rigour' of the adopted approach when carrying out evaluations. 'Rigour' in this context is considered to be a disciplined approach to the evaluation which seeks to reduce 'error' and 'bias'. There are relatively simple techniques which help avoid some obvious types of bias. For example, obtaining honest appraisals from users of a service is enhanced by ensuring participant confidentiality and anonymity. This can be difficult to achieve when interviews are used and costs of the evaluation are of concern. Nevertheless, ensuring the person who collects evaluation data is different from the person who delivered the service to the given participant is an important method of reducing bias. We consider next some aspects of 'evaluation design' with a particular focus on how to make an evaluation more rigorous.

2.3 Evaluation 'design' and outcome measures

As we have noted, the selection of an appropriate outcome measure is only one part of a service evaluation. An effective evaluation must also incorporate a 'design' which gives robust evidence of impact and change as a result of the intervention. Such an evaluation design will be able to link the intervention or service with any observed changes in outcome measures. Randomised control trials are often considered the gold standard in evaluation designs because they control for many kinds of 'bias' (see below). For this reason studies of this kind have greater weight in the research community. Other approaches are also valuable however and will often be less complex and considerably cheaper to implement. In Table 4 we list different types of evaluation design, describe their key features, and outline some limitations that should be taken into account in their use.

Table 4.
Simplified examples of evaluation design and their relative limitations

Evaluation Design	Features/example	Caution/limitations
Satisfaction survey	• Clients who have experienced a service are surveyed. • Participants asked if they were satisfied with the service. • Participants asked about the impact they believe the service had upon them. • No before-after comparison (impact of the service is based only upon participant recollection/analysis). • Potential comparison might be between sub-groups (e.g. identifying groups who were more positive than others). • Relatively cheap. • e.g. Aspinall et al (1999).	• Client 'satisfaction' is important but not a primary outcome of a service. • Descriptive / qualitative in nature (no comparison). • Recruitment bias: e.g. those who agree to be in the study may only have positive things to say. • Some participants default to positive (assuming that any intervention is better than nothing). • Pre-dispositional factors exist which mean satisfaction is linked to some participant characteristics irrespective of intervention. • See Aspinall et al (1999) for a useful overview.
Retrospective survey	• Participants who have, and have not, experienced a service are surveyed. • Comparison made between the two groups in relation to outcomes measured.	• Recruitment bias: e.g. those who agree to be in the study may only have positive things to say. • Identifying people who have not experienced a service can be problematic. • Matching problem: people who have not experienced service are typically different from people who have (e.g. may have more or less needs). • It can be difficult to be clear about the services participants have received.
Trial: Before and after design	• Clients are assessed in relation to outcome before and after intervention. • Strength is that individuals are compared to themselves. • e.g. Goodrich et al. (2006)	• No control: no comparison with people who did not experience service (i.e. would observed changes have happened anyway?).

Table 4. continued

Controlled trial: Before and after design with control	• Clients are assessed in relation to outcome before and after intervention (treatment group). • A comparison group (control) is also assessed over a similar time period. This group does not receive the service (although they may receive, e.g., a reduced or alternative service). • Strength is that individuals are compared to themselves and to the control group. • e.g. Subramanian et al (2011).	• Placement of participants in different groups can introduce bias (unless randomised – see below) • Participants usually want to be part of a treatment group rather than a control group and this is difficult to disguise in social research (in contrast to a medical trial). • Some discomfort and ethical concerns about excluding some people from treatment groups – solutions can be (a) control group have a delayed intervention, or (b) comparison is made between normal service and enhanced service. • Expensive / resource intensive.
Randomised controlled trial	• As controlled trial described above, but participants randomly allocated to 'treatment' or 'control' groups. • e.g. Reeves et al. (2004), Stelmack et al. (2008), Burggraa et al (2010).	• As controlled trial described above, but concern about group allocation bias is removed.

Some limitations are true of all evaluation designs and are emphasised in the examples provided above. For example:

- Recruitment bias. A challenge to all evaluation research of this kind is trying to account for the danger that people who choose to be part of a study may be different from those who do not. This is difficult to account for with certainty, but even in a relatively small scale satisfaction survey efforts should be made to ensure that participants have characteristics which reflect the characteristics of the population.
- Interventions need to be standardised. If the aim of the evaluation is to measure the outcome of a given service then that service must include some standardised intervention – i.e. it must be broadly the same for all the clients whether the service is a low vision therapy, mobility, information service, etc. Without this it is impossible to attribute any outcomes observed to the service.
- Appropriate outcome measures need to be used which directly align to the primary outcome under investigation.

3 Selecting appropriate outcome measures

In this section we provide a recap of key aspects that can be considered when selecting an outcome measure. We emphasise the notion of 'alignment' with a particular focus on:

- the purpose (or aims) of the service and the related information needs;
- the intended **primary outcome** of the service;
- the selected **outcome measure**; and
- the **design** of the planned evaluation.

We break down this selection process into four steps. We then present a case study based around a fictitious paediatric low vision clinic to illustrate the various stages in this process.

3.1 Step 1. Define the purpose of the service

It is useful to start by considering the primary purpose or aims of a service (for example, as articulated in a mission statement or service aims). This can then be used to help establish what types of information you are interested in collecting from the various stakeholders (i.e. what information do you need to know about in relation to the particular aim/purpose of the service?). As we have outlined in this Guide, different aims will require different types of information. This in turn will generate primary outcomes which require particular types of outcome measures. An effective evaluation will clearly align service aims, primary outcome, and outcome measure.

Another important contextual consideration of course is the requirements of the service funder. As highlighted in the introduction (and more specifically in section 1.4.4, 'Links with outcomes defined by funders'), funders of services may specify precise outcomes (and even outcome measures) that they are concerned about.

3.2 Step 2. Identify the primary outcome

As we noted in Section 1 of this Guide, outcomes can be viewed as referring to 'change' as a result of a given intervention and/or type of provision. We made a distinction between primary and ultimate outcomes (and in turn primary and ultimate outcome measures). An important step in an evaluation is to define the primary outcome. To ensure close alignment, the primary outcomes should be directly linked to the purpose or aims of the service being evaluated. If a particular primary outcome feels too broad (e.g. improved client participation, 'happiness', or 'quality of life') it may be that the focus is too much upon ultimate outcomes. Such outcomes can be considered as being broader in scope (or relate to longer term outcomes). Often such ultimate outcomes will relate to improving the 'quality of life' of the stakeholders. As we have emphasised however, quality of life is a wide concept and asking broad ranging generic questions about a person's 'quality of life' as a result of a given intervention may not be sensitive enough. Primary outcomes are more specifically linked to the activities of the service being evaluated. Identifying these primary outcomes is therefore an important step prior to selecting an outcome measure.

3.3 Step 3. Select an outcome measure aligned to the primary outcome

An important next step is to align the primary outcome with the particular outcome measure that will used to indicate change. A commonly used method to collect information is a questionnaire (usually completed through an interview). Other methods can be appropriate, depending on the particular outcome being measured.

When selecting an outcome measure it is helpful first to determine its psychometric qualities. Guidance is provided in Section 2.1 to help interpret important psychometric properties that relate to 'validity' and 'reliability'. As a minimum it will be important to establish that the instrument is appropriate for use with the particular service users (e.g. not to use a measure of functional vision designed for adults with young children). It is important to recognise however that outcome measures with good psychometric qualities which are appropriate to a particular primary outcome do not always exist. Guidance is provided in Section 2.2 to help select (or design) an outcome measure which is suitable to your needs (an outcome measure which is 'good enough'). As with the definition of the primary outcome, a useful rule of thumb when selecting

an outcome measure is to avoid being too broad. If the selected outcome measure is very generic (e.g. a broad ranging 'quality of life' measure), then this suggests a potentially poor alignment with the primary outcome.

Examples of different types of outcome measures that may be drawn upon depending on the primary outcome and the nature of provision are provided in the illustrative case studies, with a collation of example measures presented in Part 2 of this Guide.

3.4 Step 4. Planning an evaluation design

As we noted in Section 2.3, selecting an appropriate outcome measure is only part of the effective evaluation of a service. The evaluation must also incorporate a design which provides robust evidence of change as a result of the intervention – hence the term 'evaluation design'. Consideration will also need to be given to more pragmatic issues, for example: the resource available to carry out the evaluation; the time staff are able to give to the work; whether it is possible to get additional staff to support the evaluation; how much time you can reasonably expect clients to give to provide data. Examples of evaluation design and their relative limitations are listed in Section 2.3, with examples provided in the illustrative case studies in Part 2.

Sources of other information on evaluation are provided in Pocklington's briefing. An example of a project in which researchers have followed a stepped development approach similar to that described above is presented by McLinden et al (2011), (i.e. developing a pilot outcomes measure for mobility and independence specialists).

3.5 Illustrative Case Study

For this illustrative case study we introduce a fictitious Paediatric Low Vision Clinic called Ringtree. The mission statement of the clinic is:

'To empower children and young people with vision loss to help them optimise their full potential at home, school and in the wider community.'

Step 1. Consideration of the service aims and information needs

The team in the clinic has identified the following three aims that capture the scope of its current service provision:

1. To assess children's functional vision and offer practical solutions to help in the home and school environment;
2. To provide low vision aids (LVAs) to assist with everyday tasks (e.g. reading);
3. To offer information and guidance to optimise visual potential (e.g. advice on the use of the LVAs, advice on lighting).

Each aim has a different focus – the first is concerned with assessment of functional vision and ensuring the recommendations have practical significance; the second is concerned with provision of aids to help with everyday tasks; and the third relates to offering information and guidance. This articulation of succinct aims is helpful when staff in the clinic consider how best to evaluate the services they provide and outline their information needs (i.e. what information they need to know about in relation to each the aims of their service?). For the purpose of this illustrative case study, the team are particularly interested in the evaluating the second aim, namely to evaluate their provision of prescribed LVAs to determine in what ways they assist the children with their everyday tasks.

Step 2. Identifying the Primary Outcome

As a next step, the team identify their information needs and the primary outcome that will serve as the focus of the evaluation (Table 5).

Table 5.
Aligning 'service aim' and 'primary outcome'

Identified Aim of Service	Information needs (What do we want to know?)	Primary Outcome
To provide low vision aids (LVAs) to assist with everyday tasks (e.g. reading)	To find out from the children and/or their families, if there is perceived improvement in undertaking everyday tasks as a result of using a prescribed LVA.	Perceived improvement in undertaking everyday tasks as a result of using a prescribed LVA.

Step 3. Alignment of primary outcome and outcome measure

Once the primary outcome has been identified the team review potential outcome measures that can be used (see Part 2 of this Guide). They selected the Manchester Low Vision Questionnaire as it uses accessible language and is quick to administer, makes particular reference to LVA use, and has been widely used by similar services evaluating their practice. They also chose the Cardiff Visual Ability Questionnaire which had recently been designed particularly for children and is reported to have good psychometric qualities (Table 6).

Table 6.
Aligning 'primary outcome' and 'outcome measures'

Information needs	Primary Outcome	Outcome Measures
To find out from the children and/or their families, if there is perceived improvement in undertaking everyday tasks as a result of using a prescribed LVA.	Perceived improvement in undertaking everyday tasks as a result of using a prescribed LVA.	Manchester Low Vision Questionnaire Cardiff Visual Ability Questionnaire

Step 4. Planning an evaluation design

The staff then considered the evaluation design that is possible in the context of their available resource. It was decided that they would use a basic 'before' and 'after' service intervention design drawing on these outcome measures (i.e. administration of the measures before the LVA is prescribed and then again at given intervention points). The team realised that this did not control for people

improving functional vision even without the service. It was decided that if resources allow and the findings from this initial design looked promising, the team might develop a more sophisticated design in which a control is introduced (involving a parallel control of children on the waiting list). In addition, if the findings were encouraging, the team plan to consider outcome measures to record changes in relation to particular everyday tasks (e.g. objective measures of reading ability) (Table 7).

Table 7.
Choosing an appropriate evaluation design

Primary Outcome	Primary Outcome Measures	Evaluation design
Perceived change in undertaking everyday tasks as a result of using a prescribed LVA.	Manchester Low Vision Questionnaire Cardiff Children's questionnaire	A 'before and after' trial Potentially in the future a controlled trial may be designed, and considering of primary outcomes and outcome measures.

PART 2 – EXAMPLE OUTCOME MEASURES

4 Illustrative case studies of outcome measures in use

In this section we consider some examples of outcome measures, and how they have been used in evaluation design studies. The six examples differ in relation to:

- The types of service under investigation (and therefore the primary outcome of interest);
- The evaluation design;
- The psychometric qualities of the outcome measures used;
- The resources available to those carrying out the evaluation.

4.1 Case 1 – Low vision rehabilitation and children

Context

Case Study 1 is concerned with two outcomes measures relating to low vision rehabilitation. The main focus is on two outcomes measures that have been developed for children.

Case Study

Relatively few outcome measures have been developed which are specifically designed for children. One attempt to do this was the 'LV Prasad-Functional Vision Questionnaire' (LVP) which was developed in India by Gothwal et al (2003). The LVP incorporates a series of questions about the difficulty the respondent has with a series of tasks (e.g. reading a text book, reading bus numbers, lacing shoes). Hernandes Trillo (2011) carried out an investigation into the inter-relationship of low vision rehabilitation, quality of life measures and other psycho-social factors. Her research included some work with children with visual impairment and one of the outcome measures used was the LVP. This study found a correlation between children's age and LVP (i.e. older children scoring higher on the LVP) and illustrated one of the particular challenges when choosing outcome measures which are appropriate for children:

> [An] explanation for this finding is that older children are capable of performing more tasks than younger children for the simple reason of being older. The LVP questionnaire has several items including tasks that would be easier to achieve by

> older children independently of their visual impairment (e.g. "lacing shoes", "writing along a straight line", "distinguishing between a 1 pound and 2 pounds coin") and there are other tasks that very young children would never perform on their own (e.g. "walking home at night"). Hernandes Trillo (2011), p229.

Khadka et al (2010) describe the development of the Cardiff Visual Ability Questionnaire for Children (CVAQC) with the specific purpose to measure "the change in ability outcomes following interventions" (p730). The CVAQC is a 25-Item questionnaire (self completion or through interview) which has an average 10 minute administration time. It has seven subscales linked to: education (essentially access to four curriculum areas), near vision, distance vision, getting around, social interaction, entertainment, and sports. The CVAQC was subjected to a rigorous design process including item generation through focus groups, item reduction from 89 to 25 items using Rasch analysis (N=45 participants), and validity/reliability of 25-item version using Racsh analysis (N=109 participants). The authors note that sample size is limited for children under 7 years, and suggest other tools may be appropriate for that age. Some of the questions may be less relevant to primary age children (e.g. differentiation of curriculum areas language and geography; going out alone) although questionnaire scoring accounts for missing data. The difficulty of the relationship with age observed for the LVP appears to be overcome, "There was no evidence of notable differential item functioning between primary and secondary school participants" (p735), i.e. no obvious age effect. There may be some ambiguity between visual ability and independent visual ability (e.g. a child who used a low vision aid and a child who was provided with large print may report similar visual ability?) although no difficulty was picked up in the questionnaire development.

4.2 Case 2 – Eye Clinic Support Services

Context

Case Study 2 is concerned with eye clinic support services. The particular focus is on alignment between the primary outcome and the primary outcome measure.

Case Study

Patients in hospital eye clinics will receive medical diagnosis, treatment and advice by ophthalmic and optometric professionals.

In addition, patients might also receive non-medical support and advice and in some eye clinics, this type of service is sometimes formalised into an 'Eye Clinic Support Service' (ECSS). This may be staffed by, for example, nurses, rehabilitation workers, or volunteers, and sometimes the role is called an 'Eye Clinic Liaison Officer' (ECLO). There has been some interest in measuring the impact of ECSSs (e.g. Douglas, Pavey and Spurgeon, 2005).

Subramanian et al (2011) carried out a controlled trial (before and after design with control) investigating the impact of ECSSs. A simplified summary of their method is as follows. The authors recruited participants with severe sight loss attending eye clinics. Participants were interviewed when recruited and again approximately three months later. The interviews included five validated questionnaires including: The Adaptation to Vision Loss-12 (AVL-12); WHO (Five) Well-Being Index; Manchester Low Vision Questionnaire; Low Vision Quality of Life (LVQOL) questionnaire; and the MOS Social Support Survey. Between the interviews some of the participants had received support from an ECSS (intervention group, N=10) and some participants had not (non-intervention group, N=17). In spite of the authors effort to carefully design a rigorous trial they did not find that the intervention group improved more than the non-intervention group in relation to the outcome measures used.

Subramanian et al (2011) offer a number of possibilities of why there was no significant effect (e.g. small sample size, poor sensitivity of the outcome measures used, and the intervention was not intensive enough to show a difference). All these are likely to be true but of particular relevance to this Case Study is the importance of identifying the primary outcome of interest. The outcome measures used in the study had the following foci:

- The use of low vision aids (Manchester Low Vision Questionnaire);
- Mobility, activities of daily living, access to literacy, self esteem and visual functioning (Low Vision Quality of Life);
- Self esteem (AVL-12; WHO Well-being index);
- Social contact (MOS Social Support Survey).

ECSSs may have an ultimate goal of improving patient's self esteem, social contact, use of low vision aids, etc. However, it is unlikely that an ECSS will have these as primary aims. Rather a primary aim of the service is likely to relate to the referral and signposting of patients to other services. It is possible therefore that that one reason the study

did not show a significant effect was that the primary outcome of interest was not aligned to the primary outcome measure. In fact the authors noted:

> The ECLO [Eye Clinic Liaison Officer] may be able to provide a more rapid referral to social services and although no significant differences were found [..], participants from the ECLO group were more likely to have received a social services visit than patients in the non ECLO group. (p167).

An alternative design with a focus upon a primary outcome of 'speed of referral' may therefore have proved more successful in demonstrating the impact of the ECSS. Johnston et al (2011) present a list of the primary aims of ECLOs and some research findings which had greater linkage to these, particularly noting that ECLOs were associated with the following outcomes: greater clinical staff efficiency; greater clinical staff satisfaction; and quicker processing of Certificates of Visual Impairment (resulting in more expedient referrals to social care).

4.3 Case 3 – Low vision services

Context

Case Study 3 is concerned with outcome measures used by low vision rehabilitation services. This case study combines the use of outcome measures with good psychometric qualities and a rigorous and sophisticated randomised controlled trial design.

Case Study

Low vision rehabilitation services take many forms. For example, in the UK these services are often provided by optometrists working in a hospital eye clinic, and particularly focus upon providing low vision devices (Reeves at al, 2004). Stelmack et al (2008) carried out a randomised controlled trial (before and after design with control) investigating the impact of a low vision rehabilitation programme provided by the Department for Veteran Affairs in the USA. A simplified summary of their method is as follows. The authors recruited participants with moderate and severe sight loss from the programme waiting list. Participants were interviewed when recruited and again approximately four months later. The interviews included the administration of the Low Vision Visual Functioning Questionnaire (LV VFQ 48). The LV VFQ 48 measures performance in daily activities in four domains: reading; mobility; visual information processing; and visual-guided motor skills. Bins et al (2009) describe

the LV VFQ 48 as having good validity and reliability (p139).

Between the interviews participants were randomly allocated to receive either low vision rehabilitation (treatment group, N=62) or to remain on the waiting list and receive no/delayed service (control group, N=64). Statistically significant findings were found in which the treatment group improved in relation to all aspects of the outcome measure compared to the control group. It was also noted that there was a small decline in visual ability for the control group over the four months. They concluded:

> Thus, based on the large effect sizes observed for a variety of functional domains, the investigators conclude that at least 10 hours of outpatient low-vision therapy, including a home visit, is justified for patients moderately and severely impaired by low vision. Assigned homework that is reviewed by the instructor and patient is recommended to encourage patients to practice everyday tasks using low-vision devices and techniques. Furthermore, because of the small decline in functional ability over time observed in the waitlisted control group, the investigators recommend that low-vision rehabilitation services be offered as early as possible after visual impairment is diagnosed. Stelmack et al (2008), p616.

This study clearly has a number of strengths. Importantly it has a robust randomised controlled design in which there are limited opportunities for bias. However, in the context of this Guide it is the alignment of primary outcome to outcome measure which is particularly noteworthy. The LV VFQ 48 has many of the psychometric properties which were outlined earlier in relation to reliability and validity. Just as importantly, the LV VFQ 48 measured aspects of visual functioning which were relevant to the primary outcome of the service under investigation. Interestingly, Stelmack et al (2008) also included two other more generic outcome measures in their study: a measure of health-related quality of life (Medical outcomes Short Form–36, SF-36), and the Center for Epidemiologic Studies Depression Scale (CES-D). Statistically significant differences between treatment and control groups in relation to these outcome measures were not found. This would tend to support the importance of aligning primary outcome and outcome measure.

4.4 Case 4 – Employment services

Context

Case Study 4 is concerned with the evaluation of employment services. The researchers have attempted to design an outcome measure which aligns with primary outcomes of specialist visual impairment employment services.

Case Study

Employment rates amongst visually impaired people are very low in the UK (for example, the UK Vision Strategy highlights this as an area of particular concern). For this reason a number of employment services exist which are designed to support visually impaired people into employment. ENABLER is a three-year research project (started in 2010) which seeks to improve the employment opportunities of blind and partially sighted job seekers. The key aim of the project was to develop an employment assessment tool which was designed to:

- categorise clients into 'levels' which were indicative of their distance from the labour market (allowing employment services to offer programmes which were appropriate to the clients' needs by offering a formative assessment);
- measure clients progress over time (i.e. their 'distance travelled' towards employment); and therefore
- offer a summative assessment tool for supporting clients and evaluating services. Saunders et al (2012), p3.

Therefore one of the purposes of designing the assessment tool was so that it could be used as an outcome measure which would provide evidence of the impact of employment services for visually impaired people. The developers of the tool carried out an extensive design process consulting with employment professionals and visually impaired clients, and this was followed by a trial (and further re-development). To this extent the outcome measure can claim considerable face validity, although at time of writing limited analysis of the psychometric qualities of the measure have been carried out (e.g. construct validity, reliability, Rasch Analysis). Part of the design process involved reviewing existing employment assessment tools, but it was argued that:

> One of our rationales for selecting the areas of data collection in the ENABLER assessment toolkit was that the assessment tool focussed upon some of the specific areas where services

> for blind and partially sighted people are seeking to make a difference. Some of these are 'generic' (e.g. job search skills), while some are 'visual impairment-focussed' (e.g. mobility, information access). Douglas et al (2012), p84.

Therefore, the outcome measure was purposefully designed to be linked to the primary outcomes of the services it was seeking to evaluate. The authors were particularly concerned that more generic measures (e.g. wellbeing or quality of life measures) would not be sensitive enough to show any impact of the services. The resulting assessment tool included eight sections: employment activity, current job search activity, access to information, computer skills, independent travel, vision, health related issues, and target job.

4.5 Case 5 – Counselling services

Context

Case Study 5 is concerned with the evaluation of counselling services. The selected outcome measures have good psychometric qualities and the evaluation design is a relatively simple before and after trial.

Case Study

Hodge, Barr and Knox (2010) evaluated the effectiveness of emotional support and counselling (ESaC) provision within an integrated low vision service. The emotional support and counselling consisted of a number of 50 minute sessions with a counsellor. The evaluation had a number of parts but included a 'before and after' trial design involving 35 participants. To measure the impact of the service the 34-item 'Clinical Outcomes in Routine Evaluation – Outcome Measure' (CORE-OM) questionnaire was used. This outcome measure provides an overall score as well as four sub-scores related to wellbeing, problems/symptoms, life functioning and risk. The CORE-OM was administered twice, once before the intervention and again at the end of intervention. CORE-OM was chosen because it had undergone significant psychometric development:

> The CORE-OM has now been extensively used in assessing the efficacy of a range of psychological therapies with clients who have a variety of presenting problems. Data from the original pilot studies on the CORE-OM suggested it had considerable clinical face value, supportive validity and reliability, and distinguishes between clinical and non-clinical or general populations; research in the field since then has confirmed this

to be the case (see for example Connell et al., 2007; Evans et al., 2002; Gilbody et al., 2007). Hodge, Barr and Knox (2010), p22.

Also, the CORE-OM has been used in a range of studies assessing psychological wellbeing and therefore a substantial dataset of comparative outcome data exists. Practically, the CORE-OM is relatively quick to administer (5-15 minutes).

The evaluation found statistically significant and "considerable improvement in psychological well-being during the course of counselling" (p4). The evaluation was further enhanced by the additional qualitative data gathered and the comparisons made to the general population (which was possible because of the choice of outcome measure). There were some aspects of the evaluation design which weakened the strength of the findings, e.g. no control group (psychological well being might have been expected to improve over time even without the intervention), and the therapist involved in the intervention was also responsible for administering the outcome measure (participants may have felt some obligation to describe an improvement). Nevertheless, the evaluation provided useful evidence and is an example of a good alignment of primary outcome and outcome measure.

4.6 Case 6 – Intervention in relation to literacy / reading performance

Context

Case Study 6 is concerned with the evaluation of reading rehabilitation programme. The selected outcome measure has good alignment with the primary outcome of the service but does not have very sophisticated psychometric qualities. The evaluation is a relatively simple before and after trial design.

Case Study

Goodrich, Kirby, Wood and Peters (2006) evaluated the effectiveness of a reading rehabilitation programme (which involved prescribing optical reading devices, the training in the use of those devices, and training in the use of closed-circuit televisions). The evaluation included a 'before and after' trial design involving 64 participants involved in the service. To measure the impact of the service the authors developed the 'Reading Behaviour Inventory' (RBI). This outcome measure included five questions in relation to types of material read (e.g. novels, newspapers, medicine labels), reading

difficulty, frequency and satisfaction, and reading satisfaction compared to two months previously. In addition, the researchers measured reading speed and comprehension before and after intervention. The RBI had undergone very little psychometric development (the authors noted "research is needed to establish its reliability and validity" p167). Nevertheless, the authors specifically designed the RBI because they felt that existing outcome measures were too generic to evaluate their more targeted intervention.

The evaluation found statistically significant improvement between pre-training and post-training, and the improvement appeared to have been maintained two months after the intervention (as measured by follow-up administration of the RBI). There were some aspects of the evaluation design which weakened the strength of the findings. For example, there was no control group and reading behaviour and performance might have improved over time even without the intervention. Also, the rehabilitation worker involved in the intervention was also responsible for administering the outcome measure pre and post-training (participants may have felt some obligation to be positive about their reading). However, the potential bias may have been partially offset because follow-up data was collected by a researcher unfamiliar with prior results. A further source of potential bias was that the researchers reported relatively high participant drop-out between post training and follow-up (arguably these participants may have had a different experience of the service).

In spite of some of these evaluation design concerns and the relatively weak psychometric properties of the RBI outcome measure, the evaluation provided encouraging evidence that the service is having an impact upon client's reading behaviour.

5 A collation of some key outcome measures

The 'LOVSME Project' included a systematic review of low vision service outcomes (Bins et al, 2009). The review provides an appendix of outcome measures used in the studies identified in their review (p128-156). This is not an exhaustive list as the studies in the review were limited to studies which:

- fulfilled their criteria in relation to robust design; and
- were related to their definition of low vision services.

As noted by Bins et al, the variety of definitions of low vision services can be problematic, e.g. it can range from a narrow service which assesses and dispenses low vision aids (perhaps only including an optometrist) to a 'one-stop-shop' which includes a range of professionals providing a range of services (including counselling, mobility, low vision therapy). The range of functions included in low vision services was explored in the review (i.e. Table 1, p26), and while this included a range of intervention categories which might be included in a low vision service (e.g. "Training in eccentric fixation, mobility, ADLs etc." and "Assessment of psychological status /counselling") the review does not unpick the details of each category.

Bins et al (2009) present descriptions (including some details of psychometric properties) of 46 identified outcome measures. They categorise these measures under five headings:

- Objective measures (7);
- Functional measures (16);
- Vision-related quality of life (7);
- Mood/psychological (10);
- General health-related quality of life (6).

Among the objective measures, some relate to visual performance / function (e.g. near visual acuity) and others more specifically to reading skills (speed, accuracy and comprehension).

The 'other' categories are the focus of our analysis in this section and include more subjective measures generally based upon structured questionnaires (either self-completion or interview). These categories overlap to some extent. This is unsurprising given that outcome measures have variable breadth of focus. However, given our

argument that a crucial part of selecting outcome measure for a given evaluation is to identify a primary outcome, a more differentiated and precise breakdown of the outcome measures seems helpful. We draw upon Bins et al's list of outcome measures and consider each in relation to the following primary outcome areas:

- Mobility;
- Activities of daily living (ADL);
- Self-esteem (happiness, mental health);
- Literacy (reading and writing, access to information);
- Visual functioning;
- Use of LVAs;
- Social contact / participation;
- Use of Technology;
- Employment.

General health related quality of life measures are not included in this analysis given the potential for poor alignment with primary outcome areas.

Table 8. Examples of published outcome measures by primary outcome areas. Source of outcome measures Bins et al (2009) unless otherwise stated.

		1	2	3	4	5	6	7	8	9	
OUTCOME MEASURE	**SUMMARYOF MEASURE**	**Mobility**	**ADL**	**Self-esteem**	**Literacy**	**Visual functioning**	**Employment**	**Use of LVAs**	**Social contact**	**Technology**	**NOTES:**
FUNCTIONAL MEASURES											
The Daily Living Questionnaire (DLQ)	Ability and confidence in relation to ADL		Y								
Dependence level in ADL questionnaire	In/dependence with ADL tasks		Y								
Functional Assessment Questionnaire (FAQ)	In/dependence and difficulty with tasks; sense of 'loss' and motivation in relation tasks.										Primary outcome area unclear (original source could not be located)
Functional Assessment of Self-Reliance on Tasks; Clinician Rated scale (FAST-CR) and patient Self-Report scale (FAST-SR).	Ability to perform task (clinician and self report) - 11 items	Y	Y		Y						Key evaluation source: Babcock-Parziale et al (2005).
Functional Independence Measure for Blind Adults (FIMBA)	Frequency, ease and satisfaction for performing activities - 32 items	Y	Y	Y 6	Y			Y 3			Key source: Long et al (2000.) Found to be impractical and developed into the shorter VA-13.
Functional outcomes survey (VA-13)	Frequency, ease and satisfaction for performing activities - 13 items	Y 3	Y 5		Y 5						Short form of FIMBA, Key source: De l'Aune et al (2004). Not included in Bins et al (2009)
Functional Vision Status Questionnaire (FVSQ)	Difficulty associated with activities. Developed for identification. 15 items										Primary outcome area unclear (original source could not be located)

		1	2	3	4	5	6	7	8	9	
Independent Living Pre-Programme Assessment and Post-Programme assessment (ILPPA)	Ability with tasks. Interview and professional observation. 41 items.	Y	Y		Y						
Independent Living Assessment Inventory	Capacity to perform tasks. 47 items.	Y 6	Y 29		Y	Y 9					
Manchester Low Vision Questionnaire (MLVQ)	Use of LVAs, understanding of VI, satisfactions with service. 10 items							Y			
Melbourne Low Vision ADL Index (MLVAI)	Speed, accuracy and in/ dependence of performance on tasks. Clinician (16) and self-report (9). 25 items.		Y		Y						
Older Americans Resources and Services (OARS) Multidimensional Functional Status Questionnaire	Ability with/without support. Variety of broad task.		Y	Y							Broad questionnaire linked to old age generally and includes other sections (e.g, physical health).
Patient-Based Assessment of Difficulty in Mobility	Ability for independent mobility. 35 items	Y									
Perceived security in performing ADL	Self-rated performance. 28 items.	Y	Y		Y						
Reading Behaviour Inventory (RBI)	Reading behaviours, difficulty, satisfaction.				Y						
Veterans' Affairs Low Vision Visual Function Questionnaire (VA LV VFQ-48)	Ability to perform tasks. 48 items.	Y	Y		Y						
Visual Function Questionnaire (VF-14)	Difficulty in performing different vision-related activities. 14 items	Y	Y		Y						Developed to demonstrate impact of cataract; linked to demonstration of changes following surgery. Probably not sensitive enough for VI participants (e.g. includes questions in relation to driving)

		1	2	3	4	5	6	7	8	9	
VISION-RELATED QUALITY OF LIFE											
The Impact of Vision Impairment (IVI) profile	Participation in common daily experiences. 32 items.	Y	Y	Y	Y				Y		Focuses upon higher level participatory style of question. Primary outcome linked to participation?
13-item QOL measure	Dis/agreement with statements about vision-related tasks.	Y	Y		Y				Y		Broad yet shallow range of questions.
Low Vision Quality-of-life Questionnaire (LVQOL)	Relative problems with tasks/ activities	Y	Y	Y 3	Y	Y					
Measure of Functional and Psychosocial Outcomes of Blind Rehabilitation	See: 13-item QOL measure above										Appears to be the same measure as "13-item QOL measure"
NEI-VFQ (National Eye Institute Visual Function Questionnaire) 51 item	Range of question types including difficulty with tasks. 52 items.	Y	Y	Y	Y	Y			Y		Very broad range of outcomes measured. Some areas may not be relevant to VI people with permanent condition (e.g. driving)
NEI-VFQ (National Eye Institute Visual Function Questionnaire) 25 item (+appendix questions)	Range of question types including difficulty with tasks. 25 items (short form).	Y	Y	Y	Y	Y			Y		Very broad range of outcomes measured. Some areas may not be relevant to VI people with permanent condition (e.g. driving)
Vision Quality-of-life Core Measure (VCM1)	Problems caused by sight loss. 10 items.			Y							Original source could not be located. Short questionnaire seems particularly focussed upon high-level negative impact / mental health (e.g. embarrassment, anger, depression, loneliness)
Cardiff Visual Ability Questionnaire for Children (CVAQC)	Difficulty carrying out a number of tasks. 25 items.	Y			Y				Y	Y	Key source: Khadka et al (2010). Not included in Bins et al (2009)

		1	2	3	4	5	6	7	8	9	
MOOD / PSYCHOLOGICAL											
Adaptation to Age-Related Visual Loss (AVL) scale	Beliefs about own ability and perception of (age related) sight loss. 24 items.			Y							Also a 12 item short form (AVL-12).
Centre for Epidemiological Studies Depression Scale (CES-D)	Experience of symptoms of depression in the previous week. 20 items.			Y							Developed and standardised for the general population (not VI-specific)
Coopersmith self-esteem inventory	Participant rates their similarity to a series of characteristic statements. 50 Items			Y							Developed and standardised for the general population (not VI-specific). Adult and children version.
Elderly Care Research Center (ECRC) Coping Scale	Particpant rates utility of a variety of coping strategies. 22 Items			Y							Developed and standardised for the general population (not VI-specific)
Geriatric Depression Scale (GDS)	Series of statements about age related areas of worry/ concern. 15 / 30 items.			Y							Developed and standardised for the general population (not VI-specific)
Hope Scale	Series of statements about goal setting, response to situations, summative statements about approach to life. 12 items.			Y							Developed and standardised for the general population (not VI-specific)
Minnesota Multiphasic Personality Inventory (MMPI)	Statements about mental health. 43 items.			Y							Developed and standardised for the general population (not VI-specific). Complex measure (may require licensed test administrator).
Nottingham Adjustment Scale (NAS) and NAS2	Statements about adjustment to sight loss (incl. attitude, acceptance, anxiety). 55 / 47 items.			Y							Developed for visually impaired population.

		1	2	3	4	5	6	7	8	9	
Rosenberg Self-Esteem Scale	Series of positive and negative statements about self worth. 10 items.			Y							Developed and standardised for the general population (not VI-specific)
The Warwick-Edinburgh Mental Well-being Scale (WEMWBS)	14 positively worded item scale with five response categories of mental well-being			Y							Developed and standardised for the general population (not VI-specific). Tennant et al (2007).
Zung self-rating depression scale	Series of positive and negative statements in relation to symptoms of depression. 10 items.			Y							Developed and standardised for the general population (not VI-specific)

REFERENCES

Aspinall, P., Hill, A., Cotton, L. and McCloughan, C. (1999). Patient satisfaction at an evaluation measure of low vision services. *British Journal of Visual Impairment, 17*, 99-104.

Association of Directors of Social Services (ADSS) (2002). *Progress in sight: national standards of social care for visually impaired adults*. London: ADSS.

Babcock-Parziale, J., P.E. McKnight, and D.N. (2005). Head, Evaluating psychometric properties of a clinical and a self-report blind rehabilitation outcome measure. *J Rehabil Res Dev, 42(4):* p. 487-98

Big Lottery Fund (2004). Your project and its outcomes. **http://www.biglotteryfund.org.uk/project_outcomes.pdf**

Binns, A. et al (2009). *Low vision service outcomes: a systematic review*. RNIB. **http://www.rnib.org.uk/aboutus/Research/reports/earlyreach/Pages/low_vision_service_outcomes.aspx**

Burggraa, M., van Nispen R., Dankers B., and van Rens G. (2010). Effects of standard training in the use of closed-circuit televisions in visually impaired adults: design of a training protocol and a randomized controlled trial. *BMC Health Services Research, 19:62*.

Burns, S., and Cuppitt, S., (2003). *Managing outcomes a guide for homelessness organisations Charities Evaluation Services~* **http://www.lhf.org.uk/Publications/HomelessnessOutcomesGuide.pdf**

Chanfreau, J. and Cebulla, A. (2009). Educational attainment of blind and partially sighted pupils. Peterborough: RNIB. **http://www.rnib.org.uk/aboutus/research/reports/education/pages/educational_attainment.aspx**

Colenbrander, A. (2010). Assessment of functional vision and its rehabilitation. *Acta Opththalmol*. 88, 163-173.

Connell, J., Barkham, M., Stiles, W., Twigg, E., Singleton, N., Evans, O. and Miles, J. (2007). Distribution of CORE-OM scores in a general population, clinical cut-off points, and comparison with the CIS-R. *British Journal of Psychiatry, 190*, 69-74.

Corn A., and Erin J. (2010). *Foundations of Low Vision. Clinical and Functional Perspectives* (Second Edition). American Foundation for the Blind: New York

De l'Aune, M. Watson, G. Schuckers, P. and Ventimiglia, G. (2004). Clinical application of self-report, functional independence outcome measures in the DVA's Blind Rehabilitation Service. *Journal of Visual Impairment and Blindness*, 98, 4, 197-211

DfES (2002). *Quality Standards in Education Support Services for Children and Young People with Visual Impairment*. Annesley (Notts), DFES Publications.

De Boer, M., Moll, A., Henrica, C., Terwee, C., Volker-Dieben, H., and van Rens, G. (2004). Psychometric properties of vision-related quality of life questionnaires: a systematic review. *Opthal. Physiol. Opt*. 24: 257-273

Douglas, G., McCall, S., McLinden, M., Pavey, S. Ware, J. and Farrell, A. (2009). *International review of the literature of evidence of best practice models and outcomes in the education of blind and visually impaired children*. National Council for Special Education, Ireland. 192pp.
http://www.ncse.ie/

Douglas, G., Lynch, P. And Saunders, A. (2012). *ENABLER report: Six month trial of the ENABLER screening tool*. RNIB / University of Birmingham.

Douglas G, Pavey S and Spurgeon P (2005). How far is it possible to measure the impact of non-medical eye clinic support services? Occasional paper, January 2005, Number 6, Thomas Pocklington Trust

Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J. and Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180*, 51-60.

Gilbody, S., Richards, D. and Barkham, M. (2007). Diagnosing depression in primary care using self-completed instruments: a UK validation of the PHQ-9 and CORE-OM. *British Journal of General Practice, 57*, 650-652.

Goodrich, G. Kirby, J., Wood, J. and Peters, L. (2006). The Reading Behavior Inventory: An Outcome Assessment Tool. *Journal of Visual Impairment & Blindness 100, 3*, 164-168.

Gothwal, V. K., Lovie-Kitchin, J. E., et al. (2003). The Development of the LV Prasad-Functional Vision Questionnaire: A Measure of Functional Vision Performance of Visually Impaired Children. *Investigative Ophthalmology and Visual Science* 44(9): 4131-4139.

Graves, W., Moore J., and Patterson J. (1997). Foundations of Rehabilitation Counseling with Persons Who Are Blind or Visually Impaired. AFB: New York

Hernandes Trillo (2011). *The Impact of Psychosocial Factors on Adaptation and Quality of Life with Visual Impairment*. Unpublished PhD Thesis. Faculty of Life Sciences. University of Manchester.

Hodge, S., Barr, W. and Knox, P.C. (2011). Evaluation of Emotional Support and Counselling within an Integrated Low Vision Service - Final Report. Liverpool University.
http://www.rnib.org.uk/aboutus/Research/reports/support/Pages/emotional_support_evaluation.aspx

Johnston, S., Boyce, T., Galloway, D. and Simkiss, P. (2011). Cost-effective and high quality care: The evidence for Eye Clinic Liaison Officers – 'ECLOs'. Poster presented at Evidence Live 2011, 25-26 October 2011, London.
http://www.evidencelive.org/sites/default/files/posters/2011-poster-stevie-johnson.pdf

Khadka, J., Ryan, B., Margrain, T., Court, H., Woodhouse, J.M. (2010). Development of the 25-item Cardiff Visual Ability Questionnaire for Children (CVAQC). *British Journal of Ophthalmology, 94*, 730-735

Long, R., Crews, J, and Mancil, R. (2000). Creating measures of rehabilitation outcomes for people who are visually impaired: The FIMBA Project. *Journal of Visual Impairment and Blindness, 94, 5*, 292-306.

Low Vision Services Consensus Group (1999) Low vision services. Recommendations for future service delivery in the UK. RNIB: London.

Low Vision Working Group (2007). Recommended Standards for Low Vision Services: Outcomes from the Low Vision Working Group, commissioned by the Eye Care Services Steering Group. 2007.

Massof, R. and Rubin, G. (2001). Visual Function Assessment Questionnaires. *Survey of Ophthalmology.* 45, 6, 531-548.

Margolis, M., Coyne, K., Kennedy-Martin, T. Baker, T., Schein, O. and Revicki, D. (2002). Vision-specific instruments for the assessment of health-related quality of life and visual functioning: a literature review. *Pharmacoeconomics*. 20: 791-812.

McLinden, M., Pavey, S., Douglas, G. and McCall, S. (2011). Developing a pilot 'outcomes based' framework for mobility and independence specialists. Report for Guide Dogs, University of Birmingham.
http://www.birmingham.ac.uk/research/activity/education/victar/research/index.aspx

Myers, P., and Barnes, J. (2005). *Measuring Outcomes: guidance on outcome evaluation for Sure Start Local Programmes*. Institute for the Study of Children, Family and Social Issues, Birkbeck, University of London.

Ponchillia, P.E. and Ponchillia, S.V. (1996). *Foundations of rehabilitation teaching with persons who are blind or visually impaired*. New York : AFB Press.

Robson, C. (2002). Real World Research (2nd Edition). Oxford: Blackwell Publishing.

Reeves, B.C., Harper, R.A. and Russell, W.B. (2004). Enhanced low vision rehabilitation for people with age related macular degeneration: a randomised controlled trial. *British Journal of Ophthalmology*. 88(11): 1443-1449.

Saunders, A., Lynch, P. and Douglas, G. (2012). ENABLER report: *The development of a screening tool and distance travelled measure for employment services for blind and partially sighted people*. RNIB / University of Birmingham.

Steinberg, E. et al (1994). An Index of Functional Impairment in Patients With Cataract. *The Archive of Ophthalmology, 112*:630-8

Stelmack, J., et al., (2008). Outcomes of the veterans affairs low vision intervention trial (LOVIT). *Archives of Ophthalmology, 126, 5*, 608-617.

Subramanian A., Conway, M., and Gillespie-Gallery, H., (2011). The role of eye clinic liaison officers. City University London. **http://www.rnib.org.uk/aboutus/Research/reports/earlyreach/Pages/eclo_role.aspx**

Tennant, R., Hiller, L., Fishwick, R., Platt, P., Joseph, S., Weich, S., Parkinson, J., Secker, J., & Stewart-Brown, S. (2007). The Warwick-Edinburgh Mental Well-being Scale (WEMWBS): Development and UK Validation. *Health and Quality of Life Outcome, 5(63)*.

UK Vision Strategy (2008) **http://www.vision2020uk.org.uk/UKVisionstrategy/**

UK Vision Strategy (2011). *Seeing It My Way. A Universal quality and outcomes framework for blind and partially sighted people*. **www.vision2020uk.org.uk/ukvisionstrategy**

UK Vision Strategy, (2012). Consultation to refresh the UK Vision Strategy **http://www.vision2020uk.org.uk/ukvisionstrategy/**

UK Vision Strategy Implementation Plan for England (2009) **http://www.vision2020uk.org.uk/UKVisionstrategy/**

Van Dijk J. et al (2007). Health-related quality of life of child and adolescent retinoblastoma survivors in the Netherlands *Health and Quality of Life Outcomes* 2007, 5:65.

Weiner, W., Welsh, R., & Blasch, B. (Eds.) (2010). *Foundations of orientation and mobility (3rd ed., Vol. I)*. AFB: New York.