Target Monitoring and Evaluation: Measuring the Impact of Educational Psychology Interventions

Tom Connor

Institute of Education, University of London

May 2010

Acknowledgements

I would like to thank Andy Tolmie and Martin Cook, Sarah Davidson and Alastair Lidster for their help in preparing this thesis.

Most importantly, I would like to thank my mum and dad, without whose help and support I would never even have been able to get started.

Abstract

The aim of this research is to evaluate the effectiveness of a recently developed tool for measuring perceptions of the effectiveness of Educational Psychology (EP) interventions. The research project was derived following an adaptation of Goal Attainment Scaling into a revised format known as Target Monitoring and Evaluation (TME). Evidence was sought as to its utility within an EP service by investigating the reliability and validity of TME and whether or not this system could be used as a means to evaluate the efficacy of EP-led interventions in schools. Effective service delivery issues were considered by investigating the usability of TME, and evidence was sought from EP and school based colleagues with experience of using TME in order to investigate the practical, operational and commitment issues.

Within a mixed methods design the research aim was to compare the quantitative *objective utility* of TME (in which outcomes for children derived from TME were set against measurements of change from more "conventional" assessment tools) with the qualitative *perceived utility* of TME (including EP and teacher opinions of the efficacy of TME). The intention was to investigate the reliability and validity – and therefore credibility – of the TME approach by using an external point of reference and comparing *perceptions* of change measured by TME, compared to a more conventional quantitative measure of change.

The research focused specifically on clearly defined and related interventions, in order that TME measures of change could be compared with existing conventional measurement tools. These focused on a total of 24 TME cases completed for children within Key Stage 2 in mainstream primary schools. Quantitative “objective” data relating to both baseline and

outcome measures were collected using either a standardised literacy assessment or observation schedule. These were contrasted with teacher-based perceptions of baseline and change at outcome as measured by the TME process. The quantitative outcomes were investigated against the qualitative perceptions of the utility of TME via individual interviews with 10 EPs and 8 Special Educational Needs Coordinators (SENCOs) from schools who had experience of using TME. Each interview was transcribed and analysed using thematic analysis.

In the analysis, where positive progress was noted using TME, this was also usually observed using the more conventional forms of evaluation. However, there were inconsistencies in relation to the level of change in each case. The outcomes suggest TME was well regarded as a tool for assisting the process of setting up interventions and as a framework for the discussion at review. TME appeared less well regarded as an evaluative tool to measure outcomes for the EP service, and there were implications for increased support and training. From an evidence-based perspective, it may seem logical that the key element of EP evaluation ought to be based upon successful outcomes for children. However, according to the qualitative analysis, there were many bigger questions about such evaluations, for example, defining the nature of the EP's role, the difficulty in separating elements of influence, and the delivery of services through others.

Declaration

I hereby declare that, except where explicit attribution is made, the work presented in this thesis is entirely my own.

Word count (exclusive of appendices, list of reference and bibliography): 34,976

Contents

Appendices

List of tables

List of figures

1.0 Introduction

The research described in this thesis focuses on a recently developed tool for measuring perceptions of the effectiveness of Educational Psychology (EP) interventions. This opening chapter considers the role of the Educational Psychologist (EP), together with an account of the rise of evidence-based practice, and an investigation into the sorts of evaluation that EP and other related services are undertaking. Goal Attainment Scaling (Kiresuk and Sherman, 1968) will be introduced as a possible method for evaluation, from which the focus of the current study, Target, Monitoring and Evaluation (Dunsmuir, Brown, Iyadurai and Monsen, 2009), has evolved. The rationale for the current project is considered within both the national and local contexts, and an overview of the research questions and methodology is provided.

1.1 The *distinctive* role of the EP and Every Child Matters

EPs have, in the past, been employed within EP Services, as part of Local Education Authorities. The role has traditionally involved the provision of psychological advice for children and young people with Special Educational Needs (SEN) (see Love, 2009).

More recently, following the publication of the Children Act (2004), EP services have been afforded a wider role as part of Local Authority Children's Services (CSs). Over time, the involvement of EPs has increased within other services, such as Behaviour Support, Child Adolescent Mental Health and Youth Offending Teams. In addition, EP services are increasingly being 'sold' within a range of other markets, including to schools themselves (Association of Educational Psychologists [AEP], 2009). On the basis of funding

arrangements through local government there is a need to justify costs in terms of evidence and positive outcomes.

> *"Significant among those changes were the moves to integrated services and the requirement that services be judged not by what they do but against the published five desired outcomes for children and young people."*[AEP, 2009, p3]

The role of the EP is complex and varied, with a variety of activities and responsibilities undertaken towards achieving better outcomes within an educational, and increasingly societal, context for all young people. The Children's Workforce Development Council state that:

> *"EPs work in a variety of different ways to address the problems experienced by children and young people in education. They work directly with children and young people individually or in groups and with a wide range of other professionals to deliver their work."*[http://www.cwdcouncil.org.uk/educational-psychology/what]

The publication of Every Child Matters (ECM), (DfES, 2003) provides the framework for Local Authority Children's Services, including EP Services' operation and practice. Children's services are divided into multi-disciplinary teams, including representatives from Education, Health and Social Care, in order to promote the five key ECM outcomes, namely *be healthy*, *stay safe*, *enjoy and achieve*, *make a positive contribution*, and *achieve economic well-being*.

In their report written for the Department for Education and Skills, Farrell, Woods, Lewis, Rooney, Squires and O'Connor (2006) highlighted the complexities associated with establishing a clear and distinct role for EPs within CSs:

> *"Given the school and community context in which they work, and the fact that other professionals also work in these contexts, it is understandable that people might question the distinctive contribution that the EP brings."*[p99]

Many other authors, including Cameron (2006) and Wood (1998), have explored the difficulty in identifying the distinctive contribution of the EP. Cameron, for example, highlights the problems experienced by EPs in finding a role within the new service model for children and young people outlined within ECM, and raises concerns over how to demonstrate EP's core skills and knowledge within these contexts. Gersch (2004) describes how this has led to some considerable debate regarding the distinctiveness of the EP role in relation to other educational services, such as behaviour or Dyslexia specialists.

Further uncertainty is added by the increasingly competitive educational market in which EP services operate, where there are a range of professionals within any one team able to offer what may, at face value at least, appear to be similar services. Increasingly, schools receive delegated SEN funding, allowing for greater freedom to choose between services, and where cost may be an important consideration. As Farrell et al (op cit) highlight:

> *"EP time might be viewed as relatively expensive which raises a question about whether another professional might be able to undertake some of their activities more economically and with the same impact."*[p102]

The Children Act (op cit) introduced the need to report children's progress in relation to ECM outcomes. Farrell et al (op cit) demonstrate that EP services are increasingly focusing service delivery around these, a position likely to have been further emphasised in the four years since the publication of their report. Thus, the current context of multi-agency working indicates a rationale and a need for demonstrating evidence of EPs contributions to the ECM outcomes for children and young people:

> *"[EP services]... have begun to develop and evaluate their services around the five outcomes. It is likely that this work will be strengthened over the next few years with the result that services can target their work effectively and provide a reliable and accepted series of benchmarks against which to evaluate the effectiveness of their services in terms of promoting these five ECM outcomes."*[Farrell et al 2006, p101]

The implication is that it would be helpful to develop evaluative tools or processes to unpick the specific contribution of the EP, and how these contributions map onto ECM outcomes. However, while the rationale may be clear, the process of doing so is much less straightforward. Attention has been drawn to some of the reasons it has been difficult to identify the unique contribution. Farrell et al (op cit) state:

> *"School-based respondents were less certain about the relationship of EP work to meeting the five outcomes. This may reflect the fact that, compared to other responders, school staff indicated much less frequent direct experience of EP work with individuals, groups and systems as a whole. As a result their perception of the impact of EP work may be diminished."*[p10]

In this context, it is not surprising that in a report considering the evaluation of EP services, the AEP (op cit) note that:

> *"[EP] service planning and evaluation with respect to the five outcomes... is often in the early stages."*[p15]

What *is* clear is that there is a complex array of factors present in the means by which psychology and psychological skills are applied within the discipline of Educational Psychology (e.g. Monsen and Cameron, 2005). This creates difficulties in establishing an appropriate means by which to identify the distinctive contributions of EPs in meeting ECM outcomes for young people (AEP, 2009).

Demands placed on the profession at a national level impact on individual services at a local level, with a pressure to demonstrate effectiveness through the five outcomes. This point is addressed in the AEP report (2009). Here, local EP service managers are described as being, *"in the midst"* of developing means by which to address the distinctive contribution of EPs in meeting ECM outcomes (an issue also affecting the service in which the current research took place).

1.1.1 Perceptions of EP role in schools

Schools are recognised as one of the principal clients of EP services. Within the increasingly market driven educational economy, such as with the development of Academies and EP services increasingly being 'bought-back' by schools, there seems to be a need to build rapport and trust with schools, and to be sensitive to what they expect from an EP service. In previous research reporting on what makes for an effective / valued

EP, Ashton and Roberts (2006) report that it was the "*expert role*" offered by EPs that was most valued by school staff, ahead of aspects such as "*consultation*", which was reported by the EPs as being the most valuable contribution they could make to schools. The authors conclude that this may be indicative of a misconception of the EP role, attributable to a lack of awareness or recognition of what distinctive contributions the EP might make. Of perhaps greater concern within the current contexts was the finding that many school respondents felt the work traditionally carried out by the EPs could, theoretically, be undertaken by other professionals. This is particularly challenging where there are different values placed on EP work by school-based colleagues when making judgements about the effectiveness of the services offered:

> *"Data we received from schools suggest that teachers and head teachers tend not to view the success of EP work in terms of meeting the five ECM outcomes as strongly as do other agencies."* [Farrell et al, 2006, p101]

Such challenges serve to highlight the pressures that need to be overcome in order to ensure that EPs can evidence their own contributions towards positive ECM outcomes for children and young people.

1.1.2 Summary of section

It is clearly not straightforward to link EP activity with young people, their families, their schools and their communities with specific ECM outcomes. Nevertheless it is necessary to do so, despite, the acknowledged range of factors impacting and interacting with EP service delivery, and the difficulty in simply identifying what it is that the EP role covers, and how to respond positively to simple questions, such as:

"Okay then, what is it that EPs actually do?"[Wood, 1998, p12]

1.2 Factors affecting the evaluation of EP services

1.2.1 Perceptions of EP services

In the previous section, the rationale for EP services to develop methods to evaluate the impact of the services they offer was described. With increasing levels of scrutiny on service accountability and ever-greater emphasis on the need for evidence-based practice, EP services are now placing more importance on evaluation than at any time in the past (AEP, op cit).

Both the report of the DfEE (2000) and Farrell et al (op cit) highlight the range of sources used to indicate the value of EP work, including case studies, employer / client opinions and self-assessments. It is recognised in the DfES report (2006) that EP services operate within a complex and changing environment, and the *perceived* value of EPs is generally positive. However, in the literature there appears to be a reliance on a range of descriptions of positive EP contributions without necessarily any evidence of such effectiveness. For example, Stoker (2002) states:

> *"There are around 2,500 EPs in the UK... according to my estimate... we can assume that the profession has directly contributed to improved life opportunities for approximately 350,000 children and young people in any given year."*[p10]

In spite of positive perceptions of effectiveness, the Farrell et al (op cit) suggest that the holistic contribution of EPs is difficult to quantify. In the AEP report (2009) the point is raised that much research highlights the psychological aspects *used* by EPs in everyday practice, yet the *effectiveness* of the application of these processes is rarely evaluated. As such, the question is raised as to how these can really be called 'evaluations'.

Dowling and Leibowitz's (1994) review into the EP role describes the 'traditional' models of service appraisal. Examples include customer satisfaction surveys and the percentage of statutory assessment requests responded to within allocated time frames. The primary weaknesses of such approaches reside in the tendencies to measure out*puts* (the efforts produced) rather than out*comes* (the results), and to reflect on how the client groups *felt* about service delivery processes, rather than providing more useful data relating to specific outcomes for children.

In an examination of the contributions of EP practitioners to various electronic forums (such as "EPNET"), the current author found it was clear that the evaluation of psychology services has been, and remains, a difficult objective to achieve in quantifiable terms. In responding to the question, "How does your service evaluate itself and the contributions made to individual children?" there was considerable variability in terms of the methods undertaken, indicating that EP services respond in different ways to the same challenges.

1.2.2 The complex array of factors impacting on EP service evaluation

The AEP (2009) report highlights how difficult it is to link specific EP practice with ECM outcomes. One difficulty was the indirect way in which EPs might work with a young person or their family, where EP practices are based around processes which are much

less easily measured than specific observable or recognised outcomes. It is acknowledged that EP interventions take place in the real world, and as such may lack the controlled rigour present in other scientific exploration:

> *"It is apparent from the literature that distinctiveness is a difficult quality to identify unless it is an experimental or epidemiological study."*[AEP, 2009, p15]

The question of how to evaluate EP services effectively appears to be an easy one to ask, but a difficult one to answer. For example, while Baxter and Frederickson (2005) state that EPs should be demonstrating how service delivery results in improved outcomes for children, how this might be achieved is not addressed. Consequently:

> *"The 'messy' real world in which the EP works is rarely that simple or straightforward".* [AEP 2009, p18]

1.2.3 Summary of section

In their report, the AEP (op cit) state:

> *"Evaluation of the effectiveness of educational psychologists is [in] the majority of cases problematic because of the number of variables that intervene between the psychological input and the outcome."*[p18]

The literature highlights the complex factors affecting EP evaluation, for example, where the psychologist's work is *"invisible"* (e.g. AEP, 2005) when it comes to measurement. There is a logic in using existing evidence of what works when seeking given outcomes;

but linking EP activity with ECM outcomes is complicated by the *dual* meanings of "evidence-based practice".

Evidence-based practice may be seen, firstly, as a matter of using published research findings to guide one's actions; but, secondly, there is a need to demonstrate evidence for the effectiveness of the agreed actions in the current and local context. Difficulties may well arise when these two elements do not come together. This highlights the very real challenge of finding a framework, tool, or process by which to judge the value of the EP input (and the need and rationale for this current research). As Good, Simmons, and Smith (1998) state:

> *"No matter how great an intervention sounds, no matter how much it costs, no matter how much research has been published, and no matter how many criteria or belief systems it satisfies, if the intervention does not change the child's trajectory, then it is not effective for that child and a change is indicated."* [p68]

1.3 Evidence-based practice

Friedman (2005) notes that evidence is used to shape practice and service delivery, and to gain a clear basis upon which to make judgements about effectiveness. Fox (2003) highlights how evidence-based practice has become increasingly important in informing how EP services operate:

> *"Increasingly there is a belief that evidence-based guidelines on best practice are the cornerstone of providing professional service to the public."*[p93]

The concept of evidence-based practice in delivering services to children and young people within the educational sector has been described as an extension of the model of service delivery used within the medical sector and, increasingly, within other caring professions. Newman (1999), for example, states that evidence-based practice is the dominant ideology shaping the delivery of both medical and social services in both the United Kingdom and United States.

Cottrell (2002) describes how evidence-based practice has grown in popularity within Central Government at the same time that there has been a reduction in spending on public services. Thus, there is a need to target expenditure more effectively to achieve quality assurance in outcomes. Cameron (2006) discusses the government-led drive for UK education professionals to base practice on the best possible evidence of efficacy. Cameron suggests that while this has become common practice within psychology, this has been less widely adopted within the wider field of education generally.

Brann, Coleman and Luk (2001) note there is a need for evidence to be based upon the local context alongside the evidence outlined within the wider literature, since only the localised measurement of outcomes can provide a true account of intervention success. Such a principle would equally apply for individual pupils within an already localised context and, therefore, the implication is that measurement of impact has to be at the level of the individual.

There are articles in the literature suggesting that, historically, education has largely ignored the principle of adopting practice on the basis of evidence. For example, Marsh (2005) describes English-speaking countries as having developed education as a commodity, to be evaluated solely in terms of financial costs and benefits. Under such

circumstances, Marsh is concerned that many of the innovations in education have been made merely on the basis that, *"It seems like a good idea",* rather than on the basis of clear and rational evidence.

1.3.1 Summary of section

Within education there is currently a drive to apply the same aims as those within other professional disciplines, that is, to provide a system by which evidence can be rigorously applied to educational practice. Naturally, such systems should include the activities and interventions put in place following the involvement of an EP. There should be systems in place to continuously evaluate outcomes to keep services relevant in light of changing structures and client groups, and to ensure interventions are more than simply *"good ideas"*.

1.4 Evaluative frameworks

Greene, Caracelli and Graham (1989) describe the need to ensure that evaluations of educational programmes *"expand their repertoire"*[p1] including the use of both quantitative and qualitative methods. Nevertheless, the AEP (2009) highlight there is little evidence of how EP work is to be evaluated. Cline (1994), however, outlines a list of criteria for judging a service evaluation tool when investigating whether the services offered meet local needs. While each question relates directly to the forms of enquiry that should be undertaken when investigating the components for service evaluation, it is interesting to note the possible tension where these appear to be based on service level perspectives, and not relating directly to the experiences of the clients accessing the services. These are:

- *"Will it be credible to professional peers?*
- *Will it be credible to others outside the profession such as politicians, or other stakeholders?*
- *Will it support more effective service delivery?*
- *Will it support more effective management by providing information that is relevant and useful to managers?*
- *Will it facilitate an analysis of the costs and cost-effectiveness of different services?*
- *Will it be simple to operate, so that it does not divert professional or clerical staff from their main service delivery functions?*
- *Will the members of the service have a commitment to make it work?"*[Cline, 1994, p212]

Another factor concerns the need to recognise the changing contexts in which EPs are delivering psychology. While previous research may highlight the role and good practice undertaken by EPs, this does not reflect current initiatives or the emergence of integrated services affecting the profession as it currently operates.

1.4.1 Developing an appropriate evaluative framework

In a review of research on SEN support services, Gray (2001) found that within many local authorities, evaluation was still a *"relatively undeveloped area".* The review found that services were reliant on measures such as customer surveys, end of year questionnaires and activity monitoring, which were seen as providing an incomplete picture of the contributions and role of the EPs. Gray outlines a range of evaluative practice typically exemplified by support services, and which may include:

- The views of schools: Usually measured through questionnaires, and although there may be a range of issues covered, these often focus on 'satisfaction' based measures.
- The views of parents: Where there is a lack of consistency, with views often sought through informal processes or sometimes not at all.
- Other professionals: For example, investigating the views on the quality of liaison with a range of support service personnel.
- Pupils: While regarded as an essential component in evaluation, particularly with the emphasis on pupil participation in the SEN Code of Practice (2001), this is an area of enquiry that is relatively underdeveloped.
- Individual staff level: Performance assessment that is often at the local level and may focus on pupil outcomes or individual contributions to achieving service development goals.
- Monitoring processes: For example, monitoring the activity of staff across a fixed time period.

Through all these activities, Gray (op cit) highlights the significant challenges faced in *"disentangling"* the individual contributions across support services and schools, particularly where, as is the case in EP service delivery through consultation, psychologists are seeking to empower or advise rather than to assume direct responsibility for change.

1.4.2 Evaluation at a local level

At a local level, Gray (op cit) found that there was a lack of coordination across services, and an emphasis on individual discretion in the style and nature of the evaluative models used. Gray highlights several of the difficulties in developing rigorous evaluative models for support services, including the complexity of the situations encountered, and the difficulty in overcoming what might seem like *"yet another burden"* on staff time and resources. Nevertheless, the rationale behind the development of effective evaluation practices includes the need to be able to systematically review what is working, and to plan service developments accordingly.

1.4.3 Evaluative frameworks

Gray (op cit) provides a summary of different forms of evaluation that have been used to judge effectiveness, and how these operate across a range of dimensions, including:

- *Formative vs. Summative*: Evaluation either as part of an on-going process to support future development, or to draw summary conclusions. The implication is that good practice should have an element of both.
- *Internal vs. External*: Where evaluation is led by an internal, individual practitioner, or an external evaluator who may be seen as more 'independent'. Within education there has been an emerging emphasis on 'self-review'.
- *Judgement vs. Understanding*: Where evaluation may be based both on judging successes, as well as highlighting key issues or factors affecting practice.
- *Process vs. Outcomes*: Where processes may focus on aspects of quality assurance and whether or not services conform to current understandings of good

practice, or on outcomes addressing the positive or negative effects of service or practitioner activity.

- *Effectiveness vs. Cost-efficiency*: Where focusing on outcomes, judgements go beyond simple effectiveness to consider other factors, such as value for money, which may be increasingly important to demonstrate against potential 'competitors' within the education market place.

As well as adopting elements of all such practice, Gray (op cit) suggests that any system for evaluating support services, including EPs, should feature components such as:

- The achievement of the service and its individual staff should be reviewed against an appropriate range of outcomes
- Outcomes should be embedded in service goals
- Individual staff should be involved in identifying appropriate goals / outcomes, and in reviewing progress
- Meeting children's SEN should be a cooperative endeavour (e.g. methods should support a team approach to achievement)
- Evaluation should include all service staff and should involve all relevant partners / clients
- The evaluation approach should allow for a range of relevant data (including perceptions and more objective evidence)
- Evaluation should help ensure that professional development needs (at individual and service level) are identified
- Evaluation should help identify both positive and unacceptable practice
- Services should measure themselves against nationally recognised good practice standards

1.4.4 Summary of section

While it may not be possible to encompass all variables, the suggestion is it should be possible to consider the EP activity and its purpose, and to adopt measures directly relating to the activity when evaluating effectiveness of the process. In proposing an evaluative framework, the AEP (2009) suggests it may be *purpose* that is measured, not necessarily activity, and that the measure used should be that which fits best with the activity.

According to the AEP frameworks (2009), data required by external agencies such as Ofsted tend to be about outcomes and not EP activities per se. The evaluative matrix proposed by the AEP in this review indicates that the evaluation method chosen needs to be identified prior to the intervention taking place. This is up to the individual EP or service to determine. The question in relation to the current research is to what extent, and under what circumstances, might an approach such as "Target, Monitoring and Evaluation" be usefully applied as one such evaluation method.

1.5 Target, Monitoring and Evaluation

This section will present an approach which has been recently been proposed as a means to evaluate individual change in the context of EP service delivery through a consultation-based model. This section will review and summarise the literature surrounding this approach known as Target, Monitoring and Evaluation (TME), and its predecessor, Goal Attainment Scaling (GAS).

1.5.1 Introduction to TME

Dunsmuir et al (op cit) highlight the range of difficulties associated with the implementation of evidence-based practice within public services, most notably highlighting the work of Cottrell (2002), in which the difficulties for practitioners in keeping up to date with the evidence-based literature, to conduct meta-analyses, and to receive the required training in effective techniques are described. While Randomised Controlled Trials (RCTs) are generally believed to be the 'gold standard' research method (e.g. Torgerson and Torgerson, 2001), within the "real world" of EP service delivery such an approach is not always feasible. As Cline (1994) states:

> *"Recording the work of a service efficiently is a first step in quality assurance, but the major task must be to evaluate the impact of what is done."*[p218]

Dunsmuir et al (2009) propose TME as an approach by which to address such limitations, by providing an individualised process that is useful in defining measurable outcomes that demonstrate impact within CSs.

1.5.2 The origins of TME – Goal Attainment Scaling

Introduced in 1968 by Kiresak and Sherman, GAS was set up to provide a method of assessing individual outcomes in the mental health profession. GAS has since been applied in a large variety of settings, in health and social services, and, more recently, in teaching and psychology services (Roach and Elliott, 2005). GAS has also been described in terms of its usefulness within the profession of Educational Psychology (Frederickson, 2002).

Smith (1994) describes the process of GAS within a step-by-step nine-point guide, a summary of which follows.

1) *Identify the issues that will be the focus of treatment* – GAS focuses on those problems, symptoms or issues that treatment is expected to change.
2) *Translate the selected problems into at least three goals* – There are no restrictions on the sorts of goals that can be set up.
3) *Choose a brief title for each goal*
4) *Select an indicator for each goal* – The indicator is the behaviour, skill or process that most clearly represents the goal and can be used to indicate progress in meeting the goal.
5) *Specify the expected level of outcome for the goal* – The goal setter makes a prediction of the status of the client on the selected goal at the end of the intervention.
6) *Review the expected level of outcome* – Are the indicator and expected outcome consistent with the goal title?
7) *Specify the "somewhat more" and "somewhat less" than expected levels of outcome for the goal*
8) *Specify the "much more" and "much less" than expected levels of outcome* – Creating a total of five levels of outcome (progress that is much or somewhat less than expected; expected; somewhat or much more than expected)
9) *Repeat these steps for each of the three or more goals*

1.5.3 The utility of Goal Attainment Scaling

The evidence for GAS is mixed. For example, in a review of 91 studies, Cytrynbaum, Ginath, Birdwell, and Brandt, (1979) conclude that given the inherent procedural and methodological problems, the reliability and validity of GAS is deeply questionable. However, in a similar review of studies, Cardillo and Smith (1994) adopt the opposite perspective and recommend that GAS is an especially sensitive and useful measurement for evaluating change. Hurn, Kneebone, & Cropley (2006) suggest that if measures based on adaptations of GAS-based systems meet acceptable standards of reliability, validity and sensitivity, they may achieve similar levels of credibility and acceptance as a standardised and evidence-based measure of individual outcome.

In their review, Hurn et al (op cit) found evidence that, as a system, GAS has undergone many adaptations. Thus, different studies claiming to be adopting GAS-based principles may be employing systems that make direct comparisons impossible. Boothroyd and Banks (2006) additionally state that GAS does not correlate well with standardised measures. It is acknowledged (e.g. Smith, 1994) that GAS does not provide information on the level of change or adjustment in comparison to the national or local normative standards. Nevertheless, owing to its individualised approach, it may be considered unsurprising that the correlation with standardised measures is not strong, yet conversely it is precisely this aspect of GAS that acts as its primary strength.

Smith (1994) highlights the continued interest in GAS in terms of its individualised approach to measurement, and the possibility of using it *alongside* standardised diagnostic measures of outcome. Extending this point further, Smith outlines the weakness in using *"traditional"* methods to measure change, whereby a particular test or set of (often

standardised) measures are used to evaluate outcomes for all students involved within an intervention. In contrast, GAS is concerned only with those individual behaviours that the intervention is intended to change. GAS relies upon the development of outcome measures which have been specifically tailored to the individual child (or group, system, etc) whose progress is to be measured. This does, however, raise some concern that what is being measured may be quite vague or difficult to define.

Nearly 30 years ago, Evans (1981) described the positives of adopting a GAS-based approach within a mental health context, and acknowledged that these are equally likely to apply within other kinds of service. In this regard, it is the process of simply setting goals that has a positive effect on intervention outcome – in other words, where the evaluation becomes a key feature of the intervention. Nevertheless, in a more recent context, Dunsmuir et al (2009) found difficulties when implementing GAS within an EP service context. Further, Hurn et al (op cit) found that only a small number of case studies purportedly using GAS had adhered to the basic criteria, and had gone on to develop their own adaptations for pragmatic purposes. TME can be seen as one such adaptation.

1.5.4 Differences between TME and GAS

Dunsmuir et al (op cit) propose TME as a streamlined, distinct system of GAS, in order to address concerns regarding its applicability within an EP context. In terms of function, TME has been set up specifically as a tool for EPs, in order to fit within a typical consultation-driven method of service delivery.

Like GAS, TME allows the setting of up to 3 targets that can be linked directly to interventions, the principal difference being that targets can be set without the need to

define the discrete outcome levels (for example, specifically what "somewhat more than expected" is), and therefore, with a higher degree of simplicity. Dunsmuir et al (op cit) argue that TME is clearer than GAS in setting baselines, requiring only a rating assigned on a 0 – 10 Likert scale prior to and following intervention

TME appears less time intensive than GAS, therefore fitting within an EP service delivery context, where the number of visits to schools (and, by implication, time available for agreeing targets, baselines, expected outcome levels, and reviewing them with a given timespan) is limited. Nevertheless, Dunsmuir et al (op cit) claim that TME preserves the active participation of the consultee within GAS.

1.5.5 The process of TME

In an initial EP-led consultation, up to three targets are agreed, which should be "SMART" in nature (i.e. in relation to the child's needs, targets should be *Specific, Measurable, Achievable, Realistic* and *Time-bonded*). Targets are given a "*baseline*" rating using the Likert scale (1-10), which can be derived from a variety of measures, including curriculum assessment, National Curriculum levels, and p-scales. A TME form is provided by way of example in **Appendix 1**.

The baseline ratings for each target are complemented by a second rating, in which the consultee gives an "*expected*" rating on the same Likert scale. This represents where, with intervention, it is felt the child should achieve on each target within a review date set within 6 – 8 weeks of the original consultation. At the review stage, an "*outcome*" rating is allocated for the level the consultee perceives the child has achieved. The TME approach allows progress from "baseline" to "outcome" to be measured along a Likert scale. Likert

scale outcomes in relation to the "expected" progress for each target set can be compared and described under five inter-level measurements, coded as:

- Worse progress (actual rating is below that agreed at baseline)
- No progress (baseline performance is maintained)
- Some progress (outcome is rated as less than expected, but above baseline)
- Expected level of progress (actual rating matches expected rating)
- Better than expected progress (actual outcome exceeds expected rating)

One of the perceived strengths of TME is that it is designed to be specific to one individual at a time. Thus, while it will not generate standardised data, it is set up to provide an assessment of a given child's most probable response to the intervention ("expected level of outcome") which will differ from one child to another, and which may be used to augment data gathered from other sources, such as standardised or qualitative outcome indicators.

Dunsmuir et al (op cit) highlight the importance of setting clearly articulated and agreed targets in achieving the best outcomes with TME. They highlight the need for the development of coherence and continuity across professionals (whether this be in or out of the school setting) and between home and school, in both the devising and implementation of targets and intervention plans.

1.5.6 TME and consultation

According to Dunsmuir et al (op cit), TME is designed to fit with a consultation-driven method of EP service delivery. Wagner (2000) describes consultation as:

"A voluntary, collaborative, non-supervisory approach, established to aid the functioning of a system and its inter-related systems."[p11]

Dunsmuir et al (op cit) describe an oft-cited assumption that consultative working is perceived as more cost-effective and efficient than traditional referral or clinic-based models of EP service delivery. In consultation, the principle is that the consultant is there to help the consultee through a structured problem-solving process. Where TME is used as part of this consultation procedure, a measure of perceptible change is generated, but this is only possible where there is shared agreement between EP and consultee regarding what the "problem(s)" are, and how these are going to be addressed. Dunsmuir et al (op cit) suggest the creation of TME targets indicates an acknowledgement from both parties that change is possible, while the use of a Likert scale allows for change to be measured, while not necessarily seeking a "cure" or complete resolution of the problem.

There is evidence (e.g. Cameron, 2006) to suggest the importance of developing a sense of empowerment in the problem stakeholder, such as class teachers. Dunsmuir et al (op cit) suggest that TME provides opportunities for EPs to exert interpersonal influence on consultees to enhance motivation, skills and the development of creative solutions to problems which have become 'stuck'.

1.5.7 TME pilot – outcomes and evaluation

In the pilot study completed by Dunsmuir et al (2009), TME was used in consultation by Assistant EPs (AsstEPs) and main grade EPs in two services. Data from 41 cases were analysed, with a total of 96 targets assessed, at an average of 2.34 per form. The EPs

were also asked to assign an ECM category for each target, although there is no explanation as to the basis on which EPs made the choice, with the implication that this may have been based on a subjective judgement made by the EP.

In the analysis, outcome ratings were significantly higher than the baseline ratings for the target, which the researchers described as evidence that the agreed interventions had significant positive effects as measured on the Likert rating-scale. However, in the analysis there were differences in outcome depending on whether an EP or an AsstEP had been involved in the consultation.

Dunsmuir et al (op cit) suggest this could be related to the difference between the roles and experience of AsstEPs compared to EPs in setting realistic and achievable targets (where it might be argued that AsstEP targets were set too low). Nevertheless, the differences highlight a potential methodological flaw within TME, whereby, irrespective of whether the EP is fully qualified or a trainee, the outcome data achieved through TME relies purely on the consultee's perceptions or anticipations of change, and which, using a purely TME-driven approach, are difficult to quantify objectively. Thus, there appears to be some problem in defining outcomes using a system in which administration and perceptions of behaviour and change (or ECM outcome) may vary between one consultee and another.

Furthermore, the current author might speculate that differences in outcome may be dependent on a particular agenda – especially if there is some incentive or expectation (on either the part of the consultee or consultant) to demonstrate positive progress. For example, is there a risk that a consultee might deliberately seek more positive outcomes in order to validate their own part in the intervention? There may be further difficulties in the

uniformity of comparing one set of targets with others, whereby a service-wide evaluation requires that there is some level of comparability between cases irrespective of the desirability of individualised targets. It may be argued that it is easier to attribute a rate of progress in some cases than others, depending on the need of the pupil and the intervention undertaken. In addition, there may also be differences in outcome depending upon the level of pupil involvement in the setting-up and reviewing of targets since Dunsmuir et al (op cit) state:

> *"Involvement of the child is likely to contribute to improved motivation and thus outcomes."*[p51]

Although part of the TME process involves *"Setting achievable and realistic expected ratings for targets"*, Dunsmuir et al (op cit) found that there was variability in the quality of those targets investigated. While it is not defined on what criteria the targets were compared, Dunsmuir et al (op cit) report difficulties whereby some were vague, ambiguous or lacking in specificity. Consequently, the differences in experience, methodology and training between any two members of any EP service cannot be ignored as a potentially confounding variable in any analysis of outcomes using TME. Dunsmuir et al (op cit) acknowledge the difficulty that, since EP interventions tend to be broad and varied, there may be resource implications in providing sufficient training to members of an EP service in the successful application of targets.

1.5.8 Reviewing TME targets

In contrast to GAS, the TME review meeting is undertaken with the same person who set the initial baseline targets. There appear to be very good reasons to expect that the review

should always be consistently undertaken with the original problem-stakeholder. For example, not only is it likely to enhance the development of intrinsic motivation within the consultee if they are involved at both stages of the process, but there would also be a potentially significant methodological weakness if the follow-up review meeting assessed completion of targets with somebody who may have a different perception of the target, and the severity of the problem behaviour.

However, if the same consultant and consultee are involved in both baseline and review sessions, there may still be the risk of some bias, particularly if the consultee and consultant have something to gain from positive progress being noted. If TME is to be used as an effective tool, there is a need to investigate the reliability of the baseline and outcome measures.

1.5.9 Applicability of TME

Dunsmuir et al (op cit) argue that while TME will not produce standardised data, it will augment standardised and other qualitative or quantitative outcome indicators. The claim is that, since TME targets provide a clear statement of what the intervention plan is and what the intervention is set-up to accomplish, this has the advantage of avoiding the establishment of overly broad and, therefore, not very useful goals.

On a different note, Cameron (2006) describes an ideal situation within EP service delivery whereby:

"It is the nature of the relationship between input and outcome which is the focus of the interest for the psychologist and it is this quest for explanatory factors which is the raison d'être of psychological research and theory."[p94]

TME, with its range of apparent strengths and weaknesses, has been suggested by Dunsmuir et al (op cit) as a suitable tool for investigating why and how specific variables (for example, the type of difficulty, the intervention, the quality of the target set, etc) generate specific outcomes. However, there has been a lack of evidence, either positive or negative, to illustrate this in practice.

Cameron (op cit) further defines the EP role as helping to unravel problem dimensions within specific situations or contexts. Cameron (op cit) notes that in order to accomplish this, there is the need for practitioners to use an approach that is able to provide a systematic and logical analysis of the problem, without over-simplifying what might be an extremely complex real-life environment. As a method, TME *appears* to fit well with this criterion. It can be adapted to fit novel environments, administration is a joint process with problem stakeholders and by setting appropriately SMART targets it enables complex situations to be split into smaller chunks for intervention.

The use of TME raises the question of how to attribute credit for improvement in actual performance in comparison to the baseline ratings set. For example, while it may be consultation with an EP that initiates an intervention, it is unlikely the EP will be responsible for carrying through the agreed intervention. It may be argued that without the EP involvement, some interventions may not be implemented, but the question remains as to whether or not TME is measuring the general impact of EP interventions, or whether it (over)emphasises the contribution of the individual EP concerned.

While TME targets are set in collaboration with the EP, the responsibility for the implementation of the agreed strategies rests with the consultee. However, this is not atypical of the way in which school psychology services have always been delivered. Indeed, as Wagner (2000) outlines in her description of consultation, this is a process in which the agreed interventions are usually implemented by the consultee and not the EP. As Dunsmuir et al (op cit) note, the EP provides significant input via psychological knowledge to support the development of hypotheses and intervention plans. However, as stated below by Gutkin and Curtis (1999), interventions may largely be out of the EP's control:

> *"Given the indirect service delivery nature of consultation... school psychologists cannot move into either the classroom or home to provide services on a long-term ongoing basis. The ultimate responsibility for implementing interventions in schools and homes will fall to persons such as teachers and parents, that is, consultees."*[p613]

If it is accepted that the EP is unlikely to have a direct role in working with the child but works at a process level, then this must be acknowledged in any evaluation. One difficulty may be that changes might only become evident in the longer term, whereas TME suggests measuring outcomes over a relatively short time scale of 6-8 weeks.

A further issue potentially affecting the reliability and validity of TME relates to the nature of relying upon the *perceptions* of consultees to make judgments about pupil progress. Consultees may be asked to pass judgments in extremely complicated scenarios, which may be difficult to translate onto a Likert scale despite setting SMART targets. As

discussed earlier, there may be pressure on consultees to cite progress, and there are undoubtedly difficulties in expecting consultees to make wholly reliable judgments.

Nevertheless, what is clear from the discussions in the literature is that developing the 'perfect' evaluation tool for EP interventions and contributions is far from straightforward. However, it was the need for a further investigation into the usability, validity and reliability of TME that prompted this research project.

1.6 Current research rationale

It is recognised that within the EP service in which the current research took place there are already elements of evaluative practice in place. The requirement, nevertheless, is to seek a means by which to integrate evaluation with everyday practice within the service and, if possible, within the wider remit of the multi-agency CS, in order to demonstrate evidence of effective contributions towards the five outcomes. The current research, therefore, investigated the possibility of using TME to provide the framework for determining the impact of EP interventions within these parameters.

The results of Dunsmuir et al (op cit) indicate that, in the majority of cases, there was evidence for pupil progress at the review. In relation to the evaluative criteria, Dunsmuir et al (op cit) state that:

> *"The initial feedback about the TME system is promising."*[p66]

But perhaps more crucially:

"Further research will be needed in order to ascertain whether the reliability and validity reported for GAS is maintained with TME."[p67]

Based on the materials covered in this chapter, the next section will look specifically at the research questions that are central to the investigation of the utility, validity and reliability of TME, as used as a measure of outcomes relating to interventions undertaken within the current EP service.

1.7 Development of the research questions

The report of Dunsmuir et al (op cit) indicates that the methodology employed by TME was suitable for use within an EP service. However, a number of concerns are raised. For example, at follow up, how could EPs and AsstEPs be sure that progress had been made? At face value, TME appears subjective, since the baseline and follow-up measures rely on a teacher's (or other professional's) *perception* of the child, and subsequent *perceptions* of change. The potential danger is that, in 'reality', the child may not have made any meaningful change, or, alternatively, may make more change towards the desired target than the teacher is able to acknowledge.

The research project was derived following the adaptation of GAS into the revised format of TME, and evidence was sought as to its utility within an EP Service. Evidence for credibility was sought by investigating the reliability and validity of TME and whether or not this system could be used as a means to evaluate the efficacy of EP-led interventions in schools. Effective service delivery issues were considered by investigating the usability of TME, and evidence was sought from EP and school based colleagues with experience of using TME in order to investigate the practical, operational and commitment issues. In

addition, perceptions of change from the children concerned were explored through a short interview structure.

Within a mixed methods design (for example, Johnson and Onwuegbuzie, 2004) the research aim was to compare the quantitative *objective utility* of TME (in which outcomes for children derived from TME were set against measurements of change from more "conventional" assessment tools) with the qualitative *perceived utility* of TME (including EP, teacher and child opinions of the efficacy of TME). The intention was to investigate the reliability and validity – and therefore credibility – of the TME approach by using an external point of reference and comparing *perceptions* of change measured by TME, compared to a more conventional quantitative measure of change.

Quantitative tools (such as standardised assessment and structured observations) have been something of a yardstick within educational and psychological research in assessing change, and may be regarded as the cornerstone of establishing an evidence-base for practice. Nevertheless, as has been discussed, measuring change is a difficult task and such tools are not without their own limitations. For example, at an individual level, the conventional tool may not pick up the "fine grain" of differentiation or be sufficiently sensitive to subtle changes.

The conventional tools which will be introduced in Chapter 2 were chosen owing to the need to compare the TME outcomes with another client (as opposed to service) centred reference point. These more conventional tools were identified as being a reliable and non-subjective measure by which to assess change, and to create triangulation. Triangulation has been defined by Denzin (1978) as:

"The combination of methodologies in the study of the same phenomenon... The result will be a convergence of the truth about some social phenomenon."[p14]

Although TME and the more conventional quantitative measures may have had a different focus, legitimate comparisons were achieved through focusing on specific, clearly defined areas of difficulty (either 'literacy' or 'observed classroom behaviour' based) across a number of similar cases. Consideration was also given to the quality of the targets set as an indicator of positive change.

The research questions underpinning the research are outlined in the following section. In Chapter 2 the research methodology is outlined, providing a detailed account of the philosophy underpinning the research design and an explanation of how the research questions were put to test.

1.7.1 Research questions

The research was driven by the wider theme of investigating the impact and contribution of EP interventions towards the five ECM outcomes. The following provides a summary of the key research questions underlying the investigation:

1. How well does TME operate as a means of assessing the objective impact of interventions for an EP Service?
2. What are the objective outcomes when a more established and objective form of evaluation is used?
3. How does the outcome of TME compare / contrast with the outcome of more established forms of evaluation?

4. To what extent is there a relationship between the perceived utility of TME and the objectively-measured utility of TME?

2.0 Research Methodology

2.1 Introduction to research methodology

In the course of this chapter, the philosophical approach chosen for this research project will be outlined. This will include an investigation into how the use and outcomes derived from TME compare with a more conventional assessment tool, and how stakeholders understand, interpret and use the data derived from TME.

By using a mixed methods research methodology, the approach enabled an exploration of objective data relating to outcomes, and shed light on the perceptions of TME utility held by stakeholders within both the EP service and schools, including those of the children concerned. The research focused on two forms of intervention with children, relating either to Literacy or observable behaviour difficulties in the classroom. The intention was to determine whether or not TME could be used as an approach to demonstrate (positive) outcomes following individual EP interventions with young people in schools.

2.2 Research paradigm

When establishing a research methodology, Doyle, Brady and Byrne (2009) propose that the first principle is to decide upon the research paradigm. Traditionally, research has been identified as belonging within *either* a quantitative or qualitative paradigm.

Quantitative research has been described as, *"a formal, objective, systematic process in which numerical data are utilised to obtain information about the world"* (Burns and Grove, cited in Cormack, 1991, p.140). Qualitative research tends to be concerned with meaning

and how individual people experience and make sense of the world around them (Willig, 2001). Both quantitative and qualitative research methodologies have been demonstrated to have their own strengths, (e.g. Johnson and Christensen (2004), however, within psychology, there has been a general tendency to adopt a pragmatic stance; that is the principle that researchers should use the approach or mixture of approaches that work the best in a real world situation (Johnson and Onwuegbuzie, 2004). Consequently, this research employed a mixed methods research methodology, combining both quantitative and qualitative elements, owing to its suitability within the context.

2.3 Mixed methods approach

Mixed methods research is defined by Johnson, Onwuegbuzie and Turner (2007) as:

> *"Mixed methods research is the type of research in which a researcher or team of researchers combines elements of quantitative and qualitative research approaches (e.g., use of qualitative and quantitative viewpoints, data collection, analysis, inference techniques) for the broad purposes of breadth and depth of understanding and corroboration."*[p123]

Johnson and Onwuegbuzie (op cit) discuss the concept of pragmatism and mixed methods research as offering a useful position philosophically and methodologically between quantitative and qualitative approaches. Consequently, and importantly for the current research they state:

"Taking a non-purist or compatibilist or mixed position allows researchers to mix and match design components that offer the best chance of answering their specific research questions."[p15]

Johnson and Christensen (2004) outline the advantage of using mixed methods in terms of combining the strengths of both quantitative and qualitative research processes. In addition, and of particular significance to the current research, is the means by which mixed methods can answer a broader and more complex range of research questions since the researcher is not confined to a single method or approach.

Mixed methods research designs can be classified according to two dimensions, time order (whether concurrent or sequential), and paradigm emphasis (equal status versus dominant status) (Johnson and Onwuegbuzie, op cit). The current research proposes a concurrent design, in which the methods of quantitative and qualitative data collection did not necessarily lead on from each other, and in which the emphasis was on giving equal weight to the two elements.

2.4 Research design

In order to ensure consistency in the comparison between intervention types, cases for inclusion in the quantitative element of the study were restricted to TME targets relating to EP interventions for children with either literacy or low-level behaviour difficulties. The research focused specifically on clearly defined and related interventions, in order that TME measures of change could be compared with more conventional "objective" measurement tools. These focused on TME targets for children within Key Stage 2 in mainstream primary schools.

Quantitative "objective" data relating to both baseline and outcome measures were collected using either a standardised literacy assessment or observation schedule. These were contrasted with the teacher-based perceptions of baseline and change at outcome as measured by the TME process. The quantitative outcomes were investigated against the qualitative perceptions of the utility of TME via individual interviews with EPs and SENCOs from schools who had experience of using TME.

There were a total of 14 individual cases (i.e. TME forms completed at baseline and review) included within the 'literacy' condition and 10 within the 'behaviour' condition, making a total of 24 cases for inclusion in the research.

Semi-structured qualitative interviews were held with all EPs within the service who had direct experience of facilitating the TME process with school staff, totalling 10 interviews. Interviews were also held with 8 SENCOs in schools where TME had been used as part of the EP service delivery method. In addition, short interviews were held with the children for whom the TME targets had been set up, which took place at the conclusion of the follow-up assessment.

2.5 Incorporating the Every Child Matters five outcomes

In an adaptation of the TME form for the current local authority, the EP and consultee coded, at the target setting stage, the category of Every Child Matters outcome they perceived the target to belong (i.e. *be healthy*, *stay safe*, *enjoy and achieve*, *make a positive contribution*, and *achieve economic well-being*). This was then analysed to derive

the extent to which the targets evaluated change for children in relation to specific ECM criteria.

2.6 Case selection

The research undertaken in this study was designed to fit with the everyday practice of those working within the current EP service, in which TME was used as standard by EPs in the Local Authority. Target setting and follow-up of TME was completed by the relevant EP and stakeholder and not influenced by the researcher. As the research was based around a correlational design, there was no control group, though a pilot focus group and interview were undertaken to ensure the appropriateness of the semi-structured questions. A pilot of the TME forms was completed during the autumn term 2008.

It was not possible or appropriate to be prescriptive in the nature of the specific EP-derived interventions in each case. However, there was a need to ensure a reasonable level of consistency in the format of interventions in order to make comparisons across cases viable. This was achieved by considering the nature of the intervention, and by limiting the scope of the research to specifically "literacy" and "behaviour" cases, as dictated by the targets set.

2.7 Children's views

The current researcher met with each of the children for whom TME targets were set following the collection of data at the baseline stage. No specific questions were asked, and this was used as an opportunity to meet the children and to explain that the researcher would be returning to meet them again later in the term. The purpose was to

ensure that the researcher was familiar to the children at the follow-up stage at which point four questions were asked. A total of 13 interviews were completed. Responses to these questions were written down for further analysis.

2.8 The research process

The following section provides a more in-depth summary of the research process.

2.8.1 Literacy and behaviour cases

In order to ensure a meaningful service-wide comparison across cases, quantitative analysis of the TME outcomes focused specifically on two contexts. Cases for inclusion had to meet specific criteria. These were:

- Literacy: Targets focused upon difficulties specifically in word reading and in spelling.

- Behaviour: Targets focused upon forms of observed children's behaviour causing difficulties in the classroom setting.

A system was put in place whereby completion of a TME form between EP and school triggered involvement from the researcher to visit the school to collect the more conventional data. This layer of assessment was completed within one week of the TME targets being set, with a follow-up assessment within one week of the TME targets being reviewed. There was an expectation that all cases would be reviewed, with the purpose of setting outcome ratings, between 6 and 8 weeks from the initial baseline recording.

2.8.2 Context of settings

To ensure comparison across schools, the research focused only on those children currently attending mainstream primary schools. Specifically, analysis only featured cases involving children within Key Stage 2 (aged 7 - 11 years).

2.8.3 Quantitative assessment

Criteria for inclusion of cases were based on the need to ensure that the targets assessed by TME were measurable by some form of existing assessment. The following forms of assessment were completed following the initial and follow-up TME target setting sessions

2.8.4 Literacy cases

Owing to the nature of the TME targets under assessment, it was necessary to identify a more conventional measure that was sufficiently focused to cover the types of literacy target identified. Children were therefore assessed using sub-tests of the York Assessment of Reading Comprehension (YARC). The YARC was chosen because of its recent population standardisation sample, its ease of administration, and possibility for replication across the relatively short (6-8 week) time frame given its alternative sub-tests.

Baseline and Outcome TME measures were compared with outcomes derived from the YARC. Children who were targeted for Literacy-based TME targets were assessed on the following component scales:

1. Reading Accuracy: Developed to assess the decoding / accuracy of oral reading skills in primary school age children

2. Reading Comprehension: Developed to assess children's text comprehension skills (literal and inferential meaning)

3. Because many of the children involved in the sample had some degree of literacy difficulty, it was likely that their basic reading skills were not yet automatic. Consequently, the four scales from the YARC Early Reading test were also administered in order to assess key skills / knowledge underlying reading, namely:

- Letter/ Sound Knowledge: To assess alphabetic knowledge
- Early Word Reading: To assess single word reading skills, including regular and irregular words
- Sound Isolation: Assessing phoneme awareness
- Sound Deletion: Also assessing phoneme awareness

YARC outcome scores were measured for each component scale at baseline and follow-up. Raw scores were converted to an age equivalent score in months, in order to allow for a meaningful comparison across cases, and to investigate whether children's outcomes had changed to a greater extent than could be attributed to the passage of time alone. Standardised scores were not calculated owing to the short period of time between baseline and follow-up, during which some children would have passed from one age equivalent standard score range to the next, while other children would not, thus reducing the meaningfulness of the standard scores and making it more difficult to recognise the extent of the child's progress.

2.8.5 Behaviour cases

In the collection of data relating to behaviour based targets, a structured observation was undertaken by the researcher at baseline and follow-up. This was matched as far as possible to the same day of the week, same time and same classroom activity. A record of classroom activity (such as lesson, support, number of adults in the class, etc.) was taken in order to demonstrate compatibility between observations.

The classroom observation method undertaken used a time sampling method, with a record of behaviour coded at 1 minute intervals for a period of 30 minutes. Only one code could be recorded at each time point, and which related either to 'positive' or 'negative' behaviours.

Further information on the observation schedule is provided in this chapter below, and a copy is included in **Appendix 2**.

2.9 Observation schedule

The observation schedule was adapted from that included in the Psychology in Education Portfolio by Frederickson and Cameron (1999), and was selected on the basis that the codes were indicators of clearly observable features of behaviour. This satisfied the criterion of using codes based on observable behaviour, rather than codes in which the observer was required to make a subjective judgement. This was an important consideration in order to ensure the reliability of the observation schedule.

There were a total of 14 codes on which a child's behaviour could be rated. In order to keep the analysis simple, and also owing to the relatively small sample size (and, therefore, the statistically small number of times in which some individual behaviours were likely to be observed), in the analysis the observed behaviours were collapsed into two categories. These related firstly to positive behaviour (namely "actively engaged with the task", "passively engaged with the task" and "following teacher-directed instruction"), and secondly those related to negative behaviour (i.e. the rest, including, "off-task verbal" and "off-task motor"). The data were then interpreted at two levels:

1) "Overall Observation" condition: In this analysis, the total numbers of positive and negative behaviours observed were compared, with the total number of negative behaviours subtracted from the total positive behaviours observed. Across the observation period this yielded an overall observation score with a maximum possible score of 30 and a minimum possible score of -30.

2) "On-Task Behaviour" condition: In this analysis, the focus was solely on the total number of occasions the child was observed as specifically 'on task'.

2.9.1 Inter-rater reliability

In order to ensure that the observations made had a solid degree of reliability and validity, the observations of the current researcher were compared with those of another Trainee EP working within the same service. Prior to data collection, a pilot was undertaken using the observation schedule. Three separate observations were made focusing on a different child within a 'live' classroom setting. Each researcher noted behaviours observed every

minute for 30 minutes for each child and the outcomes were compared for both conditions – overall observation and on-task behaviour.

A simple 'eyeball' check indicated that the observation ratings between the two researchers were closely matched. However, in order to be sure, a further statistical analysis was undertaken using Cohen's Kappa, a method of measuring inter-rater agreement for categorical data. Cohen's Kappa ("K") was chosen owing to it being a more robust measure than simply calculating the percentage agreement between raters, as K takes into account the agreement occurring by chance.

The results are summarised in **Tables 2.1 and 2.2** below (where P_o is the percentage agreement between the two researchers, and K is the overall Cohen's Kappa score). A full breakdown of the inter-rater analysis is provided in **Appendix 3**.

Table 2.1 Overall observation scores

	Mean
P_o	.94
K	.75

Table 2.2 On-task behaviour scores

	Mean
P_o	.87
K	.73

Although one might expect a close correlation between scores when there are only two possible outcomes in each condition (either task focused / not task focused, or positive

behaviour / negative behaviour), Landis and Koch (1977) suggest that correlations above 0.81 can be described as "almost perfect agreement". The K score in both conditions was just below this threshold, with .75 achieved in the 'overall observation' condition, and .73 achieved in the 'on-task behaviour' condition. However, some care needs to be taken when applying this label, since Landis and Koch's research was based more on personal opinion than empirical evidence.

Nevertheless, the evidence from both the K score and the simple 'eyeball test' indicate that the ratings applied by the two researchers were closely matched. Therefore, the outcomes from the inter-rater reliability check provide some solid evidence for the reliability and validity of the observations made during the data collection process in the main study.

2.10 Target setting

The analysis sought to investigate the effect of the quality of targets upon TME and objective outcomes. Post-intervention, the targets featured in the analysis were rated by a sample of 12 EPs (all of whom had experience of using TME in their casework) and 8 SENCOs (who also had experience of using TME) in relation to quality. Raters were asked to judge targets on the basis of their 'SMART-ness' (specific, measurable, achievable, realistic, time-bonded) on a scale of 1 to 5 (where 1 was equivalent to 'poor', and 5 was equivalent to 'excellent').

There were different numbers of targets included on each TME form, ranging from 1 to 3. Owing to the inconsistency in target numbers affecting both within and between case variance, the mean rating for targets from the same TME form were combined, thus producing a mean target rating for each TME case. The cases were split between the

Literacy and Behaviour conditions, and the mean rating of target quality was then measured against both the TME and objective assessment outcomes.

2.11 Time effects

The research also sought to investigate whether the length of time between baseline and follow-up had any effect on the TME outcome. The number of weeks between baseline and follow-up was calculated and compared with the outcomes according to TME to investigate for any correlation.

2.12 Qualitative assessment

In the course of the data collection, two forms of qualitative data were collected. A short interview was completed with children at the TME follow-up stage, in order to investigate their own perceptions of change during the TME time period and, in particular, how this impacted on the research question concerned with the relationship between the perceived utility of TME and the objectively-measured utility of TME.

Semi-structured interviews were held with both SENCOs and EPs to elicit their experiences of using TME, and to investigate the wider research questions.

2.12.1 Interviews with children

Four structured questions were included in order to get an indication of perceptions of change on the part of the children. The questions were:

1. Tell me a bit about how you are finding school at the moment

2. Is there anything you've noticed about your learning that has changed since we last met?

3. What do you think you are best at in school?

4. Is there anything that you would like to change about your school?

A total of 13 interviews were held with pupils at the follow-up stage, 7 of whom had targets relating to the literacy condition, and 6 of whom had targets relating to the behaviour condition. Because of the young age and learning needs of many of the children involved in the analysis, it was decided to limit the interview to four questions, themed around recent experiences in school. Interviews were not held with the remaining 11 pupils, owing either to difficulties in finding a time to interview the children (for example, if play or lunchtime immediately followed on from the assessment / observation), or if, in discussion with their class teacher, this was not felt appropriate owing to the child's perceived ability to engage with the questions.

2.12.2 Interviews with EPs and SENCOs

Interviews were chosen as the means to explore the views of the EPs and SENCOs. Given the likely complexity of the information and opinions to be explored, a more structured or quantitative measurement (e.g. questionnaire) was rejected. Interviews were set up with SENCOs following the TME review phase, and only in those situations where the SENCO had been directly involved in the TME process.

In order to ensure a representative sample across the service, interviews were held with SENCOs from at least one of the schools in which each EP worked. This allowed for an

exploration of the different experiences of working with the range of EPs and how this might affect perceptions of TME, as well as generating a large enough sample to allow for a consideration of how TME was generally regarded in schools. SENCOs were only considered for interview if the school covered the Key Stage 2 sector. SENCO interviews took place in school. Unfortunately, two SENCOs declined to take part, giving a total of 8 interviews.

Each of the 10 EPs within the service who had experience of using TME in their casework were interviewed.

Semi-structured interviews are defined by Robson (2002) as:

> *"[An interview which]... has predetermined questions, but the order can be modified based upon the interviewer's perception of what seems most appropriate.'*[p270]

Semi-structured interviews were favoured because of the ease with which these could be set up compared to other forms of qualitative research methods (such as focus groups), and in view of their compatibility with a thematic process of analysis. A qualitative interview structure was devised by the current author, following a pilot focus group with staff from one school and a pilot interview with an EP colleague held earlier in the academic year. Participants were interviewed using a semi-structured schedule containing 10 questions, focusing on participants' views regarding the usability and efficacy of the TME process, such as likes and dislikes, and the perceived strengths and weaknesses.

The interview schedule was broadly split broadly between three themes. These were:

1. *Perspectives on the evaluation of EP services*

(Example interview question: "What do you think is the purpose of evaluating EP provision?")

2. *Perspectives on the applicability of TME within as an evaluative tool*

(Example interview question: Does TME provide a suitable framework to evaluate EP interventions?)

3. *Perceptions on the utility of TME*

(Example interview question: Are there any particular strengths when adopting a TME-based approach?)

Interviewees were given the opportunity to add any further comments at the end of the interview. The interview structure was the same for all participants, except where a prompt was required in order to expand on and clarify the position and comments of the interviewee. The interview structure was flexible enough to allow for responding to and following up issues raised by the interviewee that may not have been anticipated. Prompts were administered from a pre-planned script compiled from an initial pilot focus group which took place in one primary school setting and involved the SENCO, a class teacher and a teaching assistant. A copy of the interview structure is given in **Appendix 4**.

2.13 EP and SENCO interview analysis

After transcription, the interview data were analysed using a methodology based upon the 'thematic analysis' approach. Banister et al (1994) describe thematic analysis as a means by which to present interview data in relation to specific research questions. Contents are organised under thematic headings in ways that aim to be sympathetic both to elements of the research question and the preoccupations of the interviewees. Consequently, responses were compared and coded so as to highlight and classify patterns of similar incidents and responses.

2.13.1 Semi-structured interview analysis

In this section, a summary of the thematic analysis approach employed to analyse the semi-structured interviews will be provided. This section also describes how the analysis was investigated for theme validity. Reflections on the steps taken to reduce the possible confounding variable of researcher bias are considered in Chapter 6.

2.13.2 Thematic analysis

Thematic analysis involves the creation of 'codes' to fit data and the bringing together of different elements of the data to form 'themes'. Boyatzis (1998) defines a theme as being a pattern within the data that describes and organises the observations, and aiding the interpretation of aspects of the phenomenon.

Braun and Clarke (2006) describe thematic analysis as being flexible, and as having been specifically designed for use within psychological research. Their approach was used to

guide the analysis of the semi-structured interviews, and a condensed summary of this method and how it was employed in the current research is provided in **Table 2.3** below.

Table 2.3: A summary of Braun and Clarke's (2006) *"Phases of thematic analysis"*

Phase	**Description of the process**
1. Familiarizing yourself with your data	Transcribing data, reading and re-reading the data, noting down initial ideas.
2. Generating initial codes	Coding interesting features of the data in a systematic fashion across the entire data set, collating data relevant to each code.
3. Searching for themes	Collating codes into potential themes, gathering all data relevant to each potential theme.
4. Reviewing themes	Checking if the themes work in relation to the coded extracts and the entire data set generating a thematic 'map' of the analysis.
5. Defining and naming themes	Ongoing analysis to refine the specifics of each theme, and the overall story the analysis tells, generating clear definitions and names for each theme.
6. Producing the themes	Selection of vivid, compelling extracts examples, final analysis of selected extracts, relating back of the analysis to the research question and literature.

Alongside its flexibility and compatibility with the current research methodology, thematic analysis was selected on the basis that, as an approach, it does not require the researcher

to have a detailed theoretical or technological knowledge of qualitative research models (Braun and Clarke, 2006). Correspondingly, Braun and Clarke describe thematic analysis as being:

> *"A more accessible form of analysis, particularly for those early in a qualitative research career."*[p81]

2.13.3 Thematic Analysis Process

All interviews were taped using a Dictaphone (with any initially occurring thoughts and reflections of the researcher documented). Each interview was transcribed by the current researcher, with further records kept regarding the comments raised by each interviewee, and any connections among their thoughts and experiences.

Following the completion of transcriptions, individual record sheets for each interview were produced, containing mind-maps (described by Buzan (2000) as a means by which to associate ideas, words and concepts with single words or phrases) and additional comments regarding the salient issues that were beginning to emerge. Having become familiar with the data, a bottom-up process was pursued in which codes relating to the data were applied, emerging directly from the interviewee responses.

In the process of coding, each transcript was transferred into a two-column table, with the main body of text on the left hand side, and space for codes to be recorded next to the text on the right hand side. Each transcript was coded line-by-line, and also in 'chunks' of meaningful text, where the meaning or importance of a comment only became clear in reference to further responses. Coding was initially completed by hand, with the codes

transferred onto a PC using Microsoft Word, and a "thematic map" was produced (see **Appendix 5**), providing a further opportunity to consider the coding categories, and how these connected between participants.

Having completed all the initial coding of transcripts, the component elements of each code were considered and examined for consistency or overlap with other codes. This provided the opportunity to begin defining and labeling the codes, and to link these together into hierarchical groups. Some initial codes were abandoned or collapsed together at this point owing to significant overlap with others.

Once an initial hierarchical structure was defined, the emergent codes were compared and contrasted with the research questions, in order to ensure that only those codes that significantly contributed towards the initial research brief were pursued. Themes and sub-themes then emerged from the coding groups, linking and informing the data together, and also meaningfully linking back to the research questions. An example of the thematic analysis process is shown in **Table 2.4** below.

Table 2.4 Example of the thematic analysis process

1) Interview Extract (EP1) → →	**→ Initial Codes**
"I don't know. After the initial training, it doesn't seem to have been discussed, there doesn't seem to have been follow up that's been enthusiastic and positive since that point. People seemed very receptive at the initial training about it, but it doesn't seem… Perhaps it's become a chore, the process has become a chore, and while people sort of feel that it's quite a good idea and that makes sense, it's just another thing to think about, it's another form to fill in. I've found it difficult to actually do it, even though I quite like the idea."	Uncertainty Initial training session Follow-up from initial training Lack of enthusiasm for TME Positive about the idea of TME TME perceived negatively – chore Positive about the idea of TME Other priorities Difficult to complete TME Positive about the idea of TME
2) Initial Codes → →	**→ Subthemes**
• Uncertainty • Initial training session • No follow-up from initial training • Lack of enthusiasm for TME • Positive about the idea of TME • TME perceived negatively – chore • Other priorities • Difficult to complete TME	• Change and uncertainty in working practice • Value of training • Low confidence • Low confidence • Positive perceptions of TME • Negative perceptions of TME • Time and access to the EP • Desire for ease and simplicity

2.13.4 Thematic analysis reliability

In order to ascertain the trustworthiness of the analysis undertaken, the following steps were undertaken with a Trainee EP acting as 'the interrater', who was familiar with Braun and Clarke's (op cit) model of thematic analysis.

1) Definitions of all the superordinate, subordinate and sub codes used were offered by the current researcher (see Thematic Map – **Appendix 5**)
2) The current researcher and the interrater worked together looking through two interview transcripts together and all the codes were discussed. During this process it was agreed that in one instance where there was a broad link between two codes these should be collapsed together, and a further code ("Contrasting approaches in EP case work") was added.

In the discussion of the interview transcripts, other than those described above, there was broad consensus in the coding applied.

2.14 Learning checklist

A learning checklist, based on a schedule originally developed within the focus authority's SEN service, was included for completion at the baseline and follow-up phase of data collection by the current researcher.

A total of 18 behaviours were allocated a rating from 0 to 3, in both the literacy and behaviour conditions. Behaviours rated included "Sustains good eye contact with adult", "Appropriate volume of speaking voice", and "Listens well to instructions". A copy of the checklist is included in **Appendix 6.**

The ratings were equivalent to:

- 0 = Not achieved
- 1 = Achieved with support
- 2 = Achieved independently
- 3 = Secure

In the literacy condition, the rating was based on behaviours observed in the child's approach to the literacy assessment. In the behaviour condition the rating was based on the behaviours observed during the classroom observation. Thus, learning checklist scores were generated for each case at baseline and follow-up, with a maximum possible score of 54 in each the outcomes from which are outlined below.

Learning checklist scores were then compared and contrasted with the outcomes from both the TME process and from the conventional "objective" measurements of change scores.

2.15 Ethical considerations

It was important that teacher participants in the research were able to make informed judgements about whether they wished to take part, or whether parents wanted their children's TME data to be included. An initial letter was sent to schools regarding the aims and objectives of the research, a copy of which was also given to those parents where there was EP / TME involvement.

For children's TME data to be included as part of the study, parental consent was requested at the point of referral to the service. Failure to provide parental approval would not have led to any difference in terms of the service offered by the school EP. Parental consent was obtained in order for the current researcher to meet with the children.

School consent for using the TME data was also obtained. This joint consent fulfilled the requirement laid down by the Institute of Education Ethics Committee.

Securing consent from the children themselves is an area that can present considerable difficulties for conducting research in practice settings, since the precise circumstances under which consent is needed beyond that given for the 'treatment' itself are typically unclear. Within the present research, this situation would have been simplified by children being explicitly involved in the initial target setting process between school and EP (see section 4.2), since consent for the collection of data about their progress could have been obtained at that point.

In any event, there were no parents or schools who failed to provide the necessary approval. Children in the research were not identified by name or setting. At the time of the follow-up, the children were informed about the purpose of my role and given a simplified outline of the research aims. The children were asked whether they would be willing to talk to me about their recent experiences in school, and that all the information collected would be kept confidential. The children understood that they had the option of declining to take part in the discussion, but none of the children did decline. Gaining the children's consent to use the anonymised TME data was not a requirement, but had any child expressed reservations this would have been acted upon.

The interviews with EPs and SENCOs were completed on a voluntary basis, with all participants informed about the research aims. Participants were informed that the interview records were to be kept anonymous and that although their responses would be recorded, this was purely to ensure an accurate record of what was said. In addition, an agreement was made to terminate the interview at any point should the interviewee feel uncomfortable, to remove from the transcript or audio recording anything the interviewee did not wish to be noted by others, and to provide a copy of the final report should they wish to see one.

A copy of the letter sent to school head teachers at the prior to data collection is included as **Appendix 7**.

3.0 Quantitative data results

The following chapter provides a summary of the quantitative data. In a format as laid out by the research questions, the data is presented in order to explore, firstly, TME as a means of assessing the objective impact for children of interventions for EPs across the service. This results section also presents the outcomes for the same children when a more established form of evaluation is used (the YARC) and how this compares and contrasts with the TME outcomes.

In total there were 57 TME forms (featuring 157 targets) set up and reviewed across the EP service during the data collection period, equating to a mean average of 5.7 per EP in the service. Of these, 24 TME forms incorporating 46 targets were included in the analysis, across the case lists of 10 different EPs within the focus authority. Each TME form included data relating to targets and outcomes for individual children, with only one TME form per child included for the purpose of this analysis. With the targets filtered for inclusion in this analysis, this represents a proportion of 42% of the total TMEs completed.

The first analysis (section **3.1**) consists of a comparison of the outcomes as described by the TME forms, to consider the relationship between targets and ECM outcomes, and to investigate what differences there were between the applied baseline ratings, expected ratings and outcome ratings.

In section **3.2**, TME measures of change are compared at baseline and follow-up with a more conventional measurement tool. This additional layer of assessment was completed by the current researcher within one week of the TME targets being set-up, and then again within one week of the TME targets being reviewed.

Section **3.3** investigates the quality of the targets included in the sample. Ratings of target quality in both the Behaviour and Literacy conditions were compared with the outcomes according to TME, and the outcomes according to the more conventional measurement.

Section **3.4** considers the wider effects of TME. Using the Learning Checklist described in section 2.14, baseline and follow-up observation ratings as attributed by the current researcher were compared within each condition, in order to measure for any change.

A brief summary and a discussion of the key points arising are described in Section **3.5**.

3.1 TME outcomes

This section considers the relationship between targets and ECM outcomes, and investigates the relationships between the TME baseline ratings, the expected ratings and the final outcome ratings. Firstly, the data from the literacy condition will be described, followed by a summary of the data from the behaviour condition.

3.1.1 Literacy condition

Data from 14 TME forms were analysed to assess the impact of the intervention between the setting up and the review of the TME form. For each case, EPs noted up to three targets per TME form (average = 2.14), and a total of 30 targets were included in the analysis. In two cases, there was an additional target that did not relate to the "literacy" description on the TME form, and consequently these targets were discarded for the purpose of this analysis.

Examples of targets included:

- *To recognise 5 CVC words (cat, dog, hat, mum, dad)*
- *To fix the discrimination between the letters I and E in spelling tasks*

Targets were set at baseline and reviewed between 6 to 10 weeks later, during the course of the summer term in the academic year 2008-2009, as shown in **Table 3.1** below.

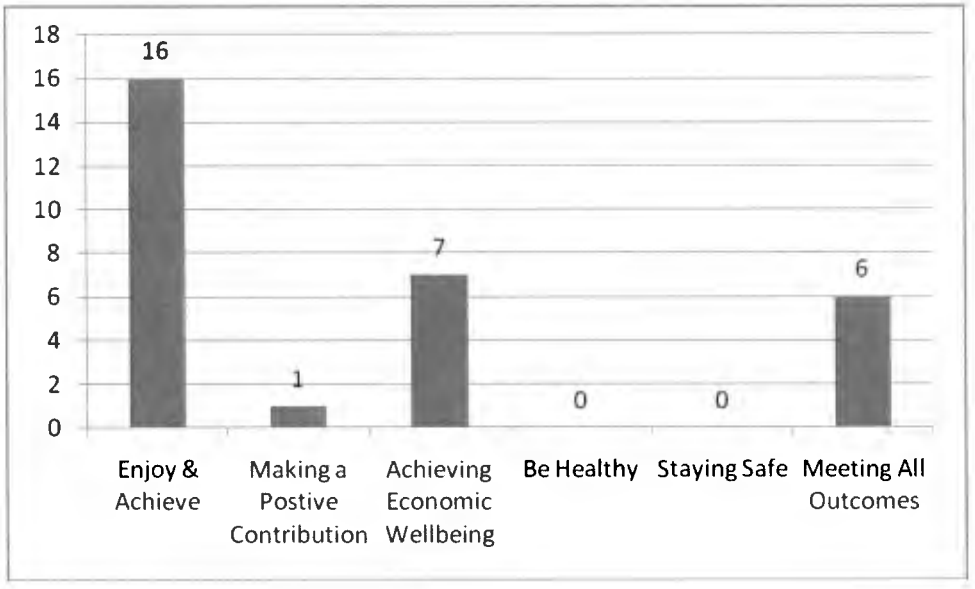
Table 3.1 TME target review timescale

	No. of targets
6 weeks	2
7 weeks	5
8 weeks	6
9 weeks	1

3.1.1.1 ECM outcomes

During the first TME session, each target was coded according to the five Every Child Matters (ECM) outcomes, to describe how the work contributed to these outcomes. Some targets were described as covering more than one ECM outcome, and for these examples, the target is recorded as meeting the range of ECM outcomes. In one case, the ECM outcome was not provided. None of the targets related to Be Healthy or Staying Safe, as might be expected, although six were described as meeting all the targets. **Figure 3.1** shows the breakdown of targets considered to meet each ECM outcome.

Figure 3.1 Literacy ECM outcomes



3.1.1.2 TME literacy outcomes

There were variable numbers of targets set for each child on their TME form (ranging from one to the maximum of three). This meant that it was not possible to reliably separate within-subject variance at the level of individual targets from that at the level of time point, since this would have required there to be a fixed number of targets. Instead, average ratings across targets were computed for each child for the baseline, expected and outcome assessments, and analyses were conducted using these averages.

Table 3.2 shows the mean ratings given for baseline, expected and actual TME outcomes for the "literacy" targets:

Table 3.2 TME baseline, expected and actual literacy outcomes

	Mean Outcome	**Standard Deviation**
Baseline	2.86	1.17
Expected	6.52	1.17
Actual	6.06	1.56

Table 3.2 indicates that the mean *actual* outcome was higher than the *baseline* rating in the literacy condition. A paired samples t-test revealed that these differences were significant ($t(14) = 6.08$, $p < 0.001$). Thus, according to the TME data, the interventions applied across the Literacy cases had a significant, positive effect on the target outcomes.

Table 3.2 also indicates that the mean *actual* outcome ratings fell short of the *expected* outcome rating. A second paired-samples t-test was applied in order to investigate the differences between the expected and actual outcome target ratings ($t(14) = 1.42$, $p = 0.179$), which indicated that there was *not* a statistically significant difference between the mean actual outcome rating and the expected outcome rating.

The lack of significant difference here may, in part, be attributed to the sizeable standard deviations, especially in relation to actual outcome score, compared to the baseline and expected scores.

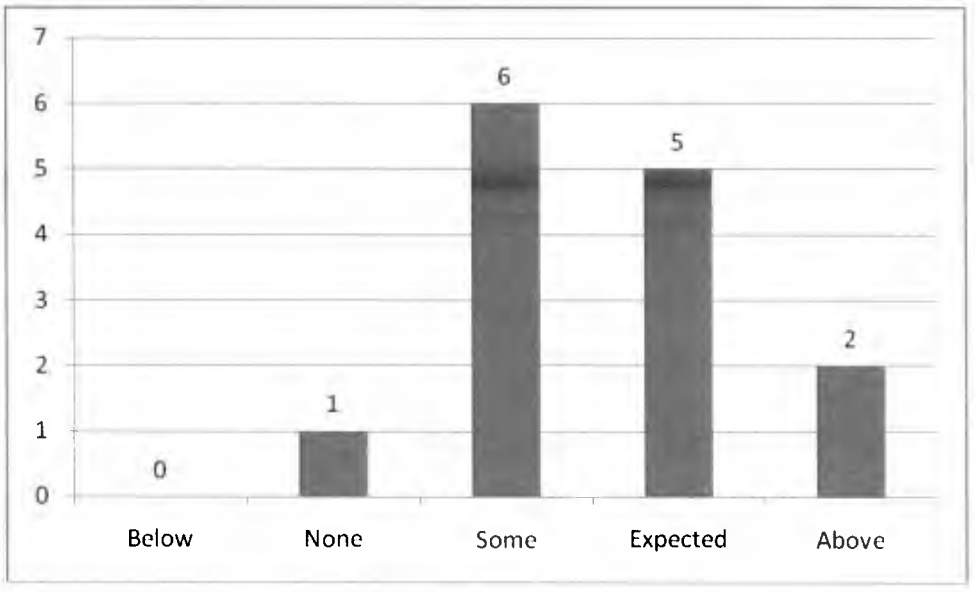
As described in the previous chapter, each actual and expected rating were compared and assigned to one of five categories. These are shown in **Table 3.3** below, including the total number of targets attributable to each category:

Table 3.3 Literacy outcomes by outcome description

Outcome Description	Total no. of cases
Worse than expected	0
No progress	1
Some progress	6
Expected progress	5
Better than expected progress	2

The data shows that of the 14 cases included, all but one case made progress. The largest individual category was that of some progress (6 cases) where progress was noted, but outcomes fell short of the expected progress level. Nevertheless, a total of 7 cases (50% of the sample) were rated as having made progress at the expected level (5 cases) or above the expected level (2 cases). **Figure 3.2** shows the proportion of outcomes in each of these categories.

Figure 3.2 Literacy outcomes by outcome description



3.1.1.3 TME literacy outcomes – time effects

In order to assess whether the number of weeks between baseline and follow-up had an impact on whether the actual outcome was above, below or at the expected level, the number of weeks were categorised as "less than 8" or "more than 8", and chi-square analyses were conducted. Results showed that there were no associations between the number of weeks between planning and follow-up, and the outcome ($p > 0.05$). Further investigation of the data, using a Spearman's Rho analysis, revealed there was no significant correlation between the number of weeks between baseline and follow-up, and the TME outcome.

3.1.2 Behaviour condition

In the behaviour condition, data from 10 TME forms were analysed to assess the impact of the intervention between the setting up and the review of the TME form. There was an average of 1.6 targets per form and a total of 16 targets were included in the analysis. In one case, there was an additional target that did not relate to the behaviour description on the TME form, and consequently this was discarded for the purpose of this analysis.

Examples of targets included:

- *To sit on the carpet tile in front of the teacher during carpet time 3 times a week for 10 minutes*
- *To keep hands to himself and not to touch other children's work*

Targets were set and reviewed within the same time frame described above, and as shown in **Table 3.4** below.

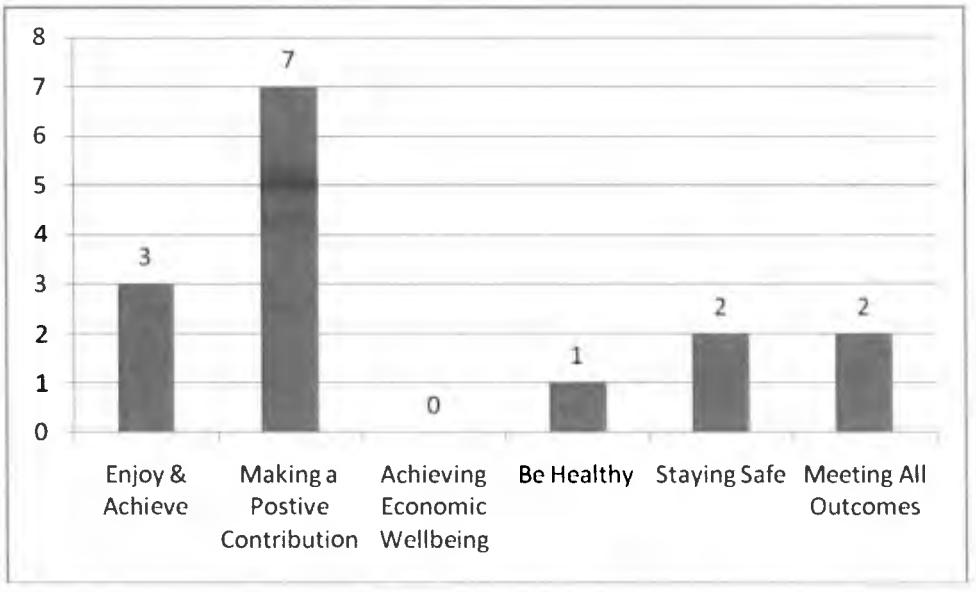
Table 3.4 TME target review timescale

	No. of targets
6 weeks	4
7 weeks	2
8 weeks	2
9 weeks	1
10 weeks	1

3.1.2.1 ECM outcomes

A similar approach was undertaken to explore how behaviour targets were considered to have met the five ECM outcomes. In this condition, ECM targets were rated as most likely to be related to Making a Positive Contribution, as shown in **Figure 3.3** below:

Figure 3.3 Behaviour ECM outcomes



3.1.2.2 TME behaviour outcomes

As described in the previous section regarding literacy outcomes, it was necessary to control for the variable number of targets per form. When a mean average was calculated this reduced the data to a total of 10 TME target ratings in the analysis.

Table 3.5 shows the mean ratings given for baseline, expected and actual TME outcomes for the "Behaviour" targets:

Table 3.5 TME baseline, expected and actual behaviour outcomes

	Mean Outcome	**Standard Deviation**
Baseline	2.50	1.11
Expected	6.10	1.17
Actual	4.90	1.49

Table 3.5 indicates that the mean *actual* outcome was higher than the *baseline* rating in the behaviour condition. A paired samples t-test revealed that these differences were significant ($t(10) = 5.62$, $p < 0.001$). Thus, according to the TME data, the interventions applied across the behaviour cases had a significant, positive effect on the target outcomes.

Table 3.5 also indicates that the mean *actual* outcome ratings fell short of the *expected* outcome rating. A second paired-samples t-test was applied in order to investigate the differences between the expected and actual outcome target ratings ($t(14) = 1.42$, $p < 0.018$). This indicated that the mean actual outcome was *significantly* lower than the expected progress rating.

As with the literacy data, the behaviour outcomes demonstrated a higher standard deviation for the actual than for the expected scores, suggesting that this may have been a common phenomenon.

Actual and expected ratings were compared, as shown in **Table 3.6** below:

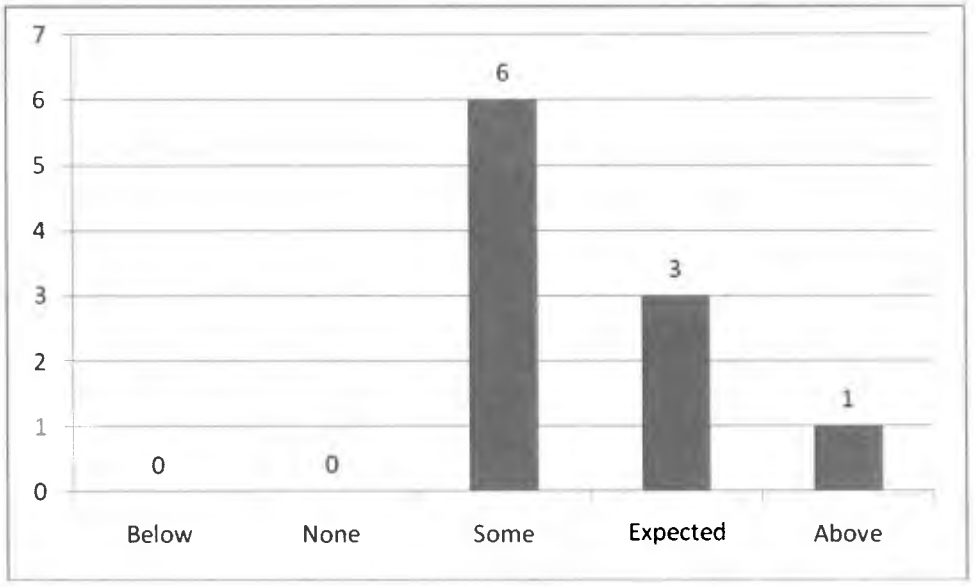
Table 3.6 Behaviour outcomes by outcome description

Outcome description	**Total no. of cases**
Worse than expected	0
No progress	0
Some progress	6
Expected progress	3
Better than expected progress	1

The data shows that in the behaviour condition, *all* the children were judged to have made at least some progress. The majority of cases were judged to be in the category of some progress (6 cases), over half those included in the sample. A smaller proportion of cases than in the literacy condition were shown to have made expected or better than expected progress (3 cases and 1 case respectively).

Figure 3.4 shows the proportion of outcomes in each of these categories.

Figure 3.4 Behaviour outcomes by outcome description



3.1.2.3 TME behaviour outcomes – time effects

As described in the previous section regarding literacy time effects, the time between baseline and follow up was investigated for impact. The number of weeks were categorised as "less than 7" or "more than 7", and chi-square analysis conducted. Results showed that there were no association between the number of weeks between baseline and follow-up, and the TME outcome ($p > 0.05$). Further investigation, using a Spearman's Rho analysis, also revealed there was no significant correlation between the number of weeks between baseline and follow-up, and the TME outcome.

3.2 TME measures of progress against conventional measurements of progress

3.2.1 Literacy cases

Baseline and outcome TME measures were compared with outcomes derived from the York Assessment of Reading Comprehension ("YARC"). The results for each YARC component scale are shown in **Table 3.7** below, including baseline and follow-up score, and the total number of cases included in each analysis.

Table 3.7 YARC outcome scores

YARC Scale	No. of cases	Mean age equivalent in months	
		Baseline	Follow-up
Reading Accuracy	13	80.7	84.2
Reading Comprehension	13	89.4	92.8
Letter/ Sound Knowledge	14	76.9	78.7
Early Word Reading	14	79.4	80.8
Sound Isolation	14	76.2	79.4
Sound Deletion	14	73.9	78.9

A full breakdown of the results indicated in Table 3.7 is included as **Appendix 8**. In the analysis, the mean follow-up score was higher than the baseline score for all six YARC component tests. Paired samples t-test revealed that these differences were significant for all components except Reading Comprehension.

For each component test, the baseline and follow-up scores were compared in order to derive an overall progress score for each case (i.e. the mean difference between the baseline and follow-up measures). The YARC progress scores were compared with the TME outcome measure, in order to investigate if there were any correlations.

No significant correlations were found except in the Sound Isolation component. Here the analysis revealed there was a significant correlation between the Sound Isolation progress measure and the TME outcome measure, $r(12) = .619$, $p < 0.05$. Therefore, higher Literacy-based TME outcomes were correlated with better performance on the sound isolation task at follow-up.

3.2.1.1 YARC outcomes time effects

Statistical analysis revealed there was no significant correlation between the number of weeks between baseline and follow-up, and any of the outcomes as measured by the YARC.

3.2.2 Behaviour cases

Baseline and follow-up TME measures were compared with outcomes derived from the classroom observations. Children's behaviour was recorded on the minute, every minute for a period of 30 minutes using the observation schedule described in Chapter 3 (and included in **Appendix 2**). There were a total of 10 children observed, with observations taken at baseline and follow-up. The data were analysed in two ways:

- The measures of positive and negative behaviours noted were compared, with the total number of negative behaviours subtracted from the total positive behaviours observed. This yielded an overall observation score with a maximum possible score of 30 and a minimum possible score of -30 ("Overall Observation" condition).

- The total number of occasions the child was observed as specifically 'on task' was separately compared ("On-Task Behaviour" condition).

The results are shown in **Table 3.8** below.

Table 3.8 Observation outcome scores

Condition	Observation	Mean score	Std. Dev.
Overall Observation Score	**Baseline**	-1.5	4.19
	Follow-up	3.8	3.94
On-Task Behaviour Score	**Baseline**	8.5	2.55
	Follow-up	11.5	4.50

3.2.2.1 Overall observation score

Table 3.8 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the overall observation condition. A paired samples t-test revealed that these differences were significant ($t(10) = 5.419$ $p < 0.01$).

The baseline and follow-up overall observation scores were compared in order to provide an overall progress score for each child (by subtracting the baseline score from the follow-

up score). Further analysis revealed there was a significant correlation between this overall observation progress score and the TME outcome measure ($r(10) = .635$, $p < 0.05$). Therefore, higher behaviour-based TME outcomes were correlated with higher overall observation scores at follow-up.

3.2.2.2 On-task observation scores

Table 3.8 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the on-task behaviour condition. A paired samples t-test revealed that these differences were significant ($t(10) = 3.402$ $p < 0.05$).

The baseline and follow-up on-task behaviour observation scores were compared, this time to provide an overall on-task behaviour progress score for each child. Further analysis revealed there was no significant correlation between the overall on-task observation progress scores and the TME outcome measure.

3.2.2.3 Behaviour observation time effects

Statistical analysis revealed there was no significant correlation between the number of weeks between baseline and follow-up, and any of the outcomes as measured by the behaviour observations (either "Observation" Progress or "On-task behaviour" Progress).

3.3 Target quality

In the analysis, targets were split into two categories, literacy and behaviour, and rated by a sample of EPs and SENCOs according to how "SMART" they felt the targets were. Targets were rated along a five point scale in which 1 represented an "excellent target", through to 5 which represented a "poor target". A mean rating was then calculated for each target, for both the EP and the school staff ratings.

As described in section **3.1.1.2**, in order to reduce the effect of within / between case variance, the mean rating for targets from the same TME form were combined, thus producing a mean target rating for each TME case.

3.3.1 Literacy targets

After controlling for the between / within case variance, there were a total of 14 target ratings analysed in the literacy condition. Overall mean ratings are shown in **Table 3.9**.

Table 3.9 Mean ratings for literacy targets

	Mean Rating	**Standard Deviation**
EP	2.7	.32
SENCO	2.4	.31

Table 3.9 indicates that the mean EP target ratings were higher than the mean SENCO ratings. An independent samples t-test revealed that these differences were significant ($t(14) = 2.803$ $p < 0.01$).

Further analysis revealed there was a significant correlation between the EP and SENCO literacy target ratings ($r(14) = .749$, $p < 0.01$). Therefore, although the absolute levels were different, higher EP ratings of target quality were likely to be matched with higher SENCO ratings of target quality, implying a difference in calibration between the two groups.

3.3.1.1 Literacy targets and TME outcomes

The ratings of target quality for the literacy targets were compared against the TME measures of outcome for each case. There were no statistically significant correlations between TME outcomes and either the EP or SENCO ratings of literacy target quality.

3.3.1.2 Literacy targets and YARC outcomes

The ratings of target quality for the literacy targets were compared against the 6 YARC measures of outcome for each case. It had been hypothesised that targets with a higher rating of quality would result in better YARC outcomes. There were no statistically significant correlations between the EP and SENCO ratings of quality against YARC measures of reading accuracy, reading comprehension, early word reading, sound isolation or sound deletion.

Analysis revealed that there was a significant correlation between EP ratings of target quality and the YARC outcome measure on letter sound knowledge ($r(14) = .601$, $p < 0.05$). Therefore, higher EP ratings of target quality were correlated with better performance on the YARC letter sound knowledge task at follow-up. There was no statistically significant relationship between the SENCO target ratings and the letter sound knowledge performance.

Interestingly, the correlation shown here between EP ratings and letter sound knowledge is not the same variable on which YARC and TME outcomes were correlated (which was sound isolation). This, and the fact that it was just one index on which the significant correlations were found, suggests that these might have been random / chance results.

3.3.2 Behaviour targets

After controlling for the between / within case variance, there were a total of 10 target ratings analysed in the behaviour condition. Overall mean ratings are shown in **Table 3.10**.

Table 3.10 Mean ratings for behaviour outcomes

	Mean Rating	**Standard deviation**
EP	2.3	.34
SENCO	2.7	.51

Table 3.10 indicates that the mean EP target ratings were higher than the mean SENCO rating. An independent samples t-test revealed that these differences were not significant ($t(10) = 2.023$ $p > 0.05$).

Further analysis revealed there was no correlation between the EP and SENCO behaviour target ratings ($r(10) = .207$, $p > 0.05$). Since there was no relationship between the EP and SENCO ratings of behaviour target quality, this indicates that this was more than just a difference in relative scaling (as was the case in the literacy cases), but that there were differences in perceptions between the two groups in what was regarded as a measure of quality in a behaviour-based target.

3.3.2.1 Behaviour targets and TME outcomes

The ratings of target quality for the behaviour targets were compared against the TME measures of outcome for each case. It had been hypothesised that targets with a higher rating of quality would result in better TME outcomes. No support was found for this hypothesis – there were no statistically significant correlations between TME outcomes and either the EP or SENCO ratings of behaviour target quality.

3.3.2.2 Behaviour targets and observation outcomes

The ratings of target quality for the behaviour targets were compared against the two observation measures of progress ("Overall observation" progress and "On-task behaviour" progress).

There were no statistically significant correlations between the overall observation scores and the EP ratings of target quality. However, analysis revealed that there was a significant correlation between SENCO ratings of target quality and the on-task behaviour progress score ($r(10) = .712$, $p < 0.01$). Therefore, higher SENCO ratings of target quality were correlated with higher scores of on-task behaviour progress in the observation at follow-up. There was no statistically significant relationship between the SENCO target ratings and the Observation progress outcome score.

Although there is the same risk as with the literacy target correlation described in the previous section (where SENCO target ratings correlated with letter sound knowledge) that the SENCO target rating / on task behaviour outcome score correlation may have been

attributed to chance, it is interesting to note this relationship. It indicates that EPs and SENCOs are basing judgements on target quality on different aspects. SENCOs appear to give higher value to those targets which are more explicit in their expectations for children to work independently and on-task, as opposed to those targets which may apply to behaviour in more general terms, as shown by the lack of relationship between SENCO target ratings and the overall behaviour progress scores.

3.4 Learning checklist

As described in section 3.14, an additional "Learning Checklist" score was generated at baseline and follow-up. A total of 18 behaviours were allocated a rating by the current researcher from 0 to 3, in both conditions. Behaviours rated included "Sustains good eye contact with adult", "Appropriate volume of speaking voice", and "Listens well to instructions". These ratings were equivalent to:

- 0 = Not achieved
- 1 = Achieved with support
- 2 = Achieved independently
- 3 = Secure

A total of two learning checklist scores were generated for each case at baseline and follow-up, with a maximum possible score of 54, the outcomes from which are outlined below.

3.4.1 Learning checklist – literacy condition

A total of 8 cases in the literacy condition were included in the analysis. The learning checklist outcomes for these cases are shown in **Table 3.11** below:

Table 3.11 Learning checklist literacy ratings

	Mean Rating	**Standard deviation**
Baseline	20.4	6.9
Follow-up	23.5	6.5

In the literacy condition, Table 3.11 indicates that the mean learning checklist rating at follow-up was higher than the mean learning checklist rating at baseline. A paired samples t-test revealed that these differences were significant ($t(8) = 1.74$ $p < 0.01$).

3.4.2 Learning checklist – behaviour condition

A total of 6 cases in the behaviour condition were included in the analysis. The learning checklist outcomes for these cases are shown in **Table 3.12** below:

Table 3.12 Learning checklist behaviour ratings

	Mean Rating	**Std. Dev**
Baseline	12.3	5.5
Follow-up	14.8	6.4

Table 3.12 indicates that the mean learning checklist rating at follow-up was higher than the mean learning checklist rating at baseline. A paired samples t-test revealed that these differences were not significant ($t(6) = 2.71$ $p > 0.05$).

3.4.3 Summary of key points arising from the learning checklist

The learning checklist provides some opportunity to compare directly the outcomes in both conditions. The purpose of the checklist was to investigate beyond the targets identified in the TME process themselves, in order to identify whether there were any further positive outcomes derived from the TME process and the interventions identified therein.

The analysis shows that in both conditions there was an increase in the checklist score at follow-up, with the results indicating a trend towards further benefits beyond the progress made in relation to the targets themselves. Furthermore, in the literacy condition, the difference between the baseline and follow-up rating was statistically significant. Therefore, it could be argued that there are a range of additional benefits to the child which are not being measured by TME and which may, ordinarily, not be noted as part of any wider evaluation of children's progress using TME. This relates directly to the comments raised by both EPs and SENCOs described in the next chapter.

It should be noted that the numbers of cases involved in the analysis of the learning checklist was very small (8 and 6 in the literacy and behaviour conditions respectively), thus these results should be treated with some degree of caution owing to the difficulty in generalising from such a small sample.

The increased outcome ratings may have been attributed to the research methodology itself. For example, the learning checklist was a very general measure, and there may have been a degree of subjectivity in the application of a rating (between 0 and 3) for each behaviour observed. Owing to time limitations before the start of data collection, there was no opportunity to compare the inter-rater reliability of the measure by comparing outcomes derived from the current researcher with the outcomes from another observer.

3.5 Summary of points arising from the quantitative results

TME ratings are, by their very nature, subjective. Different raters may use different means and criteria to determine both baseline and expected "scores" and offer different ratings of the same performance in a given child. This might also be reflected by the different perceptions of target quality (see **section 3.3**).

However, it can be argued that TME ratings, with their associated discussions and observations, provide reinforcement of the *consultation approach*... i.e. the shared involvement of teachers, classroom assistants, SEN coordinators, other professionals, and parents in planning and monitoring and decision-making; and such ratings may also provide at least a partial basis for determining the value of the particular contribution of the EP.

Further, whatever the means adopted to produce the baseline and follow-up ratings of a pupil's learning, what matters is the extent of change between the ratings produced at baseline and at follow-up as an indicator of progress (or lack of progress). These ratings may be based upon difficult-to-quantify features, or subject to variations beyond the remit /

scope of TME, as much as any change in the specific and targeted functioning (see **section 3.1**).

Formal assessment on normative test material may be insufficiently sensitive, especially over a short time scale, to produce differences in quantitative scores despite some positive progress. TME ratings can only provide an overall impression of child performance by means of ratings chosen from a 1-10 point scale. "Traditional" measures are quantitative and specific concerning, for example, sight vocabulary acquisition, or reading comprehension, and the data available involve a less constrained range. Therefore, one might well predict *some* positive association between the two sets of scores relating to a group of children, but not a close correlation, since one is not linking like with like (see **section 3.2.1**). A note of caution is implicit in the findings, to avoid the possible risk highlighted by Feuer, Holland, Green, Bertenthal, and Hemphill (1999) whereby:

> *"We note that it is often possible to calculate arithmetic linkages that create misleading interpretations of student performance."*[pp4]

Thus, TME, as a tool specifically cited for its individual approach, is unlikely to match perfectly with any alternative form of assessment (such as assessment, test or observation) from either a teacher's or researcher's perspective.

On the other hand, when it comes to *behaviour*, any form of assessment is likely to involve a greater degree of subjectivity (and, presumably, a reflection of individual tolerance thresholds). Therefore, it is not surprising that TME ratings and scores from more traditional assessment tools are more likely to show a clear correlation (see **section 3.2.2**).

TME ratings and traditional scores may be seen as complementary, and the relative emphasis upon one or the other will vary according to the precise nature of the situational demand and the purpose to which the information is to be put. TME ratings may well provide a general indicator that progress is taking place and, by implication, a general measure of the usefulness of the shared consultations and of the coordinating role of the EP. Therefore, in TME it appears that a veneer of quantitativeness is superimposed upon what is a largely qualitative exercise. In addition, TME appears to provide a structure within which to direct attention towards specific, realistic, and individualised targets as part of the means of determining progress, as reflected within the 'expected' rating.

4.0 Qualitative results 1: Children's views and perspectives

At the 'follow-up' stage, the researcher met again with 13 children within the school setting for whom TME data had been collected to complete a short interview. The aim was to learn more about the experiences in school following the setting of TME-based targets from the children's perspectives, and how these perceptions might affect both the TME and the more conventionally measured outcomes.

Owing to the age of the children involved in the study, the interview was limited to four questions, presented in a semi-structured form, and designed to fit into a general discussion with the researcher. These were:

1. Tell me a bit about how you are finding school at the moment
2. Is there anything you've noticed about your learning that has changed since we last met?
3. What do you think you are best at in school?
4. Is there anything that you would like to change about your school?

The main findings are presented below. Excerpts from the interview material are used to illustrate the emergent themes.

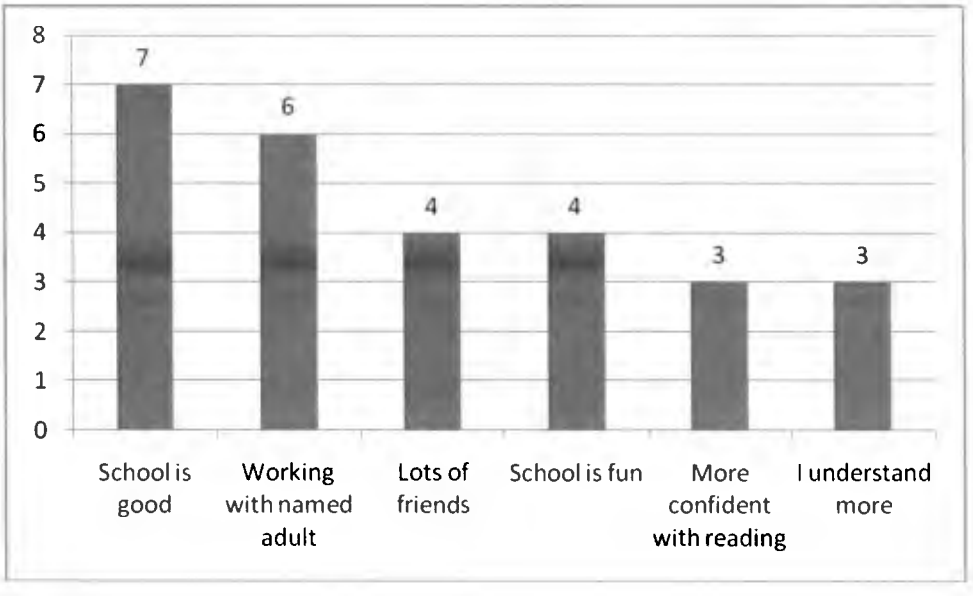
4.1 Current experiences of school

The majority of the pupils interviewed held positive views regarding their current experiences within school. A breakdown of the most frequently occurring responses is shown below in **Figure 4.1**.

The most positive elements of school related to the quality of their relationships with adult teaching and support staff, and to the support received within the classroom. Specifically, children (n = 6) reported that since first meeting with the researcher, they had been working closely with a named adult either in or outside the class.

> *"I really like working with Mrs X in the mornings and also with Mrs Y. She helps me to understand what I have to do."*(Year 4, male, literacy condition)

Figure 4.1 Summary of positive experiences of school



One pupil commented that she had missed quite a lot of school that term and was concerned about the impact this was going to have on her learning. It appeared that the targeted support she had received had helped overcome some of this anxiety:

> *"I was ill earlier in the term and I was worried I'd missed loads of stuff, but it's been okay because I was helped to catch-up."*(Year 6, female, literacy condition)

There was, however, an indication that some of the children interviewed were becoming over-reliant on adult support staff and that support given outside the classroom was not generalising into the classroom. For example:

> *"I'm enjoying school, but it's hard to focus sometimes when I'm in the class. I like it more out of the class and can work more quietly with Mrs X."*(Year 5, male, behaviour condition)

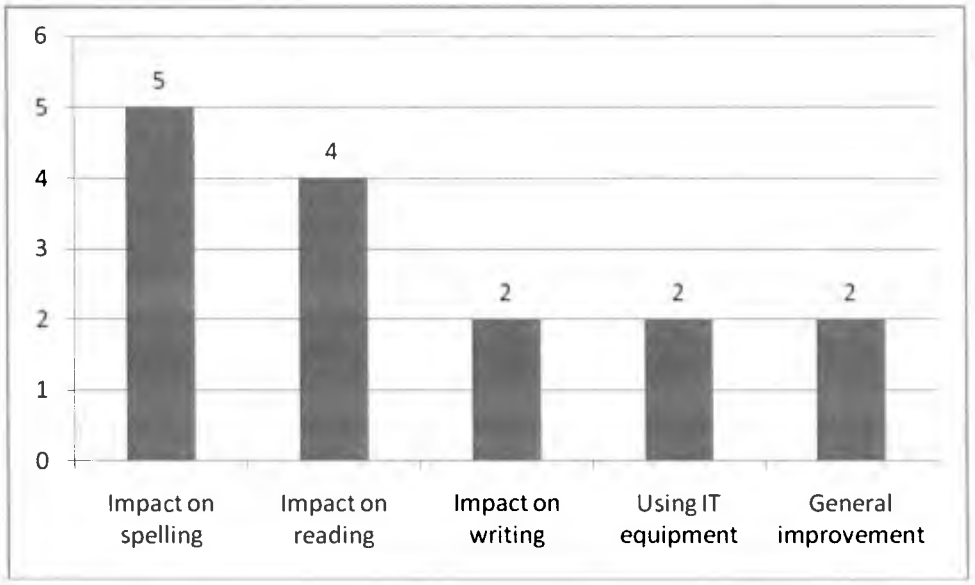
4.2 Differences perceived between baseline and follow-up

Children were asked to comment on anything they had noticed about their learning or school that was different since the current researcher had met them at previously. A summary of the most frequently appearing responses is provided in **Figure 5.2** below.

Children across *both* conditions reported an improvement in their spelling skills (literacy condition n = 3; behaviour condition n = 2) and in their reading skills (both conditions n = 2). Although the numbers of pupils are small, this suggests that the nature of the support offered led to increased feelings of confidence with literacy, irrespective of the type of target set:

"I've been finding my spellings a bit easier, and I've been in a group with Mrs X and some other people in my class. It's been helping me with my writing and with my spelling." (Year 5, male, behaviour condition)

Figure 4.2 Perceived differences between baseline and follow-up



Despite the positive comments expressed towards school experiences, no children commented on the TME targets their teaching staff had set on their behalf. One child was able to report that he had been working towards a set of targets, although he was not able to provide any further indication of what these were:

"We've been doing more Literacy work together and now I can read some more words. I got more stickers because I did well on my targets!"(Year 4, male, literacy condition)

One child reflected on the recent SATs exams that he had completed during the period of data collection, and the negative impact this had on his desire to be in school and the impact on his learning at this time.

> *"Well, school is so much better now that the SATs are over... I'm enjoying some of my lessons again now, especially maths, and we get to do proper games lessons again."*(Year 6, male, literacy condition)

4.3 Reflections on change in school

Across the two groups, a total of seven children reported that in the future they would like to have more time spent with either their class teacher or a preferred member of support staff. These sorts of relationships appeared to be valued by the children, indicating an important factor in the progress made with their learning.

> *"I really like working with Mrs X, especially when we do fun stuff and it helps me learn and remember. I'd like to do more work with Mrs X."*(Year 3, male, literacy condition)

Nevertheless, this raises concerns with regard to the termination of support through TME. For example, while the relationships with key staff appear to be important in terms of raising confidence and feelings of being supported, there were also emerging themes relating to withdrawal and over-reliance. It may be that the termination of targeted support, which is likely to be an issue whatever the evaluation system in place, is an event that could be targeted directly through the TME planning and review phases to reduce associated difficulties.

"I know I'm leaving here soon, but I wish that I could take Mrs X with me."(Year 6, female, literacy condition)

4.4 Summary of the views and perspectives of the children

The pupil perspective, represented by this admittedly small sample, is positive, with possible implications for the perceived value and supportiveness of the intervention within which the EP is likely to have been a significant contributor and coordinator. Nevertheless, it is important to consider that there was no control or baseline in order to inform whether this was associated with the period of TME-based intervention.

Of particular significance is the lack of awareness amongst the pupils of their TME targets, and the consequent difficulties, therefore, for the children being able to make judgements themselves about their learning (or other) progress in school. Furthermore, where one pupil commented on the SATs exams, this indicates the competing demands and pressures on teaching staff time and expectations that may have impacted upon the implementation of the agreed TME targets and interventions. These themes are explored further in the next section, regarding the views and experiences of the teachers and EPs using TME.

5.0 Qualitative Results 2: EP and SENCO perspectives

Chapter 5 will investigate the outcomes from the semi-structured interview phase of the research. The qualitative outcomes are compared and contrasted in Chapter 6 with the results of the quantitative analysis reported in Chapter 4 and the analysis of the children's comments reported in Chapter 5.

5.1 Introduction

In total, the data consisted of 18 semi-structured interviews, 10 with EPs and 8 with primary school SENCOs, all of whom had direct experience of using TME. Comments received from participants in the initial pilot focus group are also reported.

5.2 Superordinate themes

In the analysis, three superordinate themes were identified relating to the use of TME within school, EP interactions, and the complexity of identifying a suitable system by which to evaluate EP impact. These were:

1) *Professional roles*: The difficulties associated with ascertaining the role of the EP across different contexts.
2) *Resistance*: The impact on motivation and engagement associated with the introduction of a new process, such as TME.
3) *The practicalities of TME*: The strengths, difficulties and applicability of TME as a means by which to evaluate EP work.

Within each of these superordinate themes, further subordinate themes and sub-themes were identified. These are presented using tables, indicating the number of EP or SENCO respondents for each sub-theme. In the main body of text, direct interview quotes are used to illustrate the data, attributed either to the EP or SENCO group. The emergent themes are discussed in relation to the original research questions in the next chapter (Chapter 7) and are described graphically in the thematic map presented below in **Figure 5.1**. The themes underlying each superordinate and subtheme are presented in **Appendix 9,** including a summary of the recurrence of each theme by EP or SENCO.

Figure 5.1 TME superordinate and subordinate themes

5.3 Superordinate theme 1: Professional roles

> *"There's rarely a true definition of what the EP role is... It would make it much easier to evaluate work if you did."*(EP5)

This theme concerns the context of professional roles and identity. It picks up on the theme of TME as an appropriate tool within the context of the EP role, and how this links with the perceived risk of using TME across the range of school and professional staff working towards possibly conflicting goals and outcomes. The perceived challenge of involving parents and children within an evaluative context is also discussed.

The data are organised in three main categories:

- The role of the EP
- The presence of other factors
- Involving parents and children

5.3.1 Role of the EP

It was noted that a number of references were made to uncertainties about the precise roles and responsibilities of the EP (with the possibility that such uncertainty will have been increased by the emergence or development of other advisory and consultative roles). The need to be clear about both why a child is being referred to the EP service and what the specific role of the EP is going to be were themes that drew a number of comments from both interview groups. There was an acknowledgement that without such shared agreement from the outset, there are likely to be situations where professionals are not

working in partnership towards the same outcomes, perceived as the source of some possible tension.

> *"We see EPs as scientific practitioners. Schools and the LA probably do not. The gap is where the uncertainties lie, very different agendas."* (EP5)

One implication is the potential usefulness of producing a guideline setting out the various services that EPs can and do provide (perhaps with illustrations from casework and other activities). A similar exercise on the part of other professionals (such as advisory teachers, outreach staff, etc) would highlight a degree of overlap but should also highlight core roles thus ensuring that demands and expectations upon EPs, and other professionals, are informed and reasonable.

It appeared from the responses obtained, depending on the nature of the individual piece of work under discussion, without shared, negotiated roles there might be very different perceptions about what constitutes a successful intervention with implications, therefore, for how that piece of work should be evaluated.

> *"Whatever that initial referral was, you'd be looking to judge against that"*(SENCO4)

> *"I'd say it's a mixed experience, but a lot of it comes down to what the definition of the EP role is and if that's not agreed [the TME] it's not useful."*(EP6)

There was recognition that progress may not be directly measured by TME if this were not linked to the original target. Thus, the need for clarity and agreement in the original EP referral was seen as important in avoiding confusion about the EP and school working

together towards the same, shared outcomes, and to avoid the possible confounding variable of evaluations being based on different perceptions of what the EP or school were going to do, or what the desired outcome might look like.

> *"But different people have different ways of defining an intervention, which may be quite confusing."*(EP9)

> *"It's hard because the EP doesn't always understand the situation in which we are working in the school."*(SENCO8)

Although some respondents saw TME as a useful means to ensure clarity about the role and referral, one EP highlighted the TME process as feeling "*somewhat back-to-front*", whereby the evaluation (i.e. TME) was perceived as "*driving*" the intervention, the intervention, therefore, having to be designed to fit the TME criteria:

> *"Interventions should be designed around the individual situation, and not be constrained by the evaluation procedure itself."*(EP1)

In a finding echoing that of Ashton and Roberts (2006), the definition of the EP role formed an additional layer of complexity, with several SENCOs reporting that they were not always clear what it was that an EP could offer. This was complicated further by a feeling among EPs that schools were not always willing to accept an alternative or different approach to casework, perhaps where they were not aware of the range or breadth of potential EP working. Thus, simply initiating a conversation on TME in some settings was perceived as challenging.

"The role seems to have changed such a lot recently, it's always changing. I'm not really sure what it is EPs do anymore".(SENCO7)

"It's not always easy to set things up."(EP8)

Furthermore, there was an acknowledgement among EPs that there may be differences in the approaches used by different EPs in working towards the same sorts of outcomes for young people.

"We've probably had a conversation before about the lack of a framework for applied psychology in educational psychology."(EP5)

Nevertheless, a unifying theme among EPs and SENCOs was that EP-led interventions should be based on the best available evidence. Responses highlighted the need to draw evidence of positive outcomes from a wide range of sources, but with implications for the time and resources available.

"We have to be able to prove that what we are doing is worth the funding, and I suppose it is the same for EPs."(SENCO6)

"We want to evaluate the provision in general to see whether we're doing the job we should be doing."(EP4)

5.3.2 The presence of other factors

The high number of possible factors underlying progress and the successful achievement of targets was identified by interviewees in both samples. Where children were judged to have made the agreed (or better) levels of progress, the difficulty in identifying the factors that made the specific difference was highlighted.

> *"I think it's very difficult to know what we add in that situation, because there are so many other factors".*(EP4)

> *"It might be the system at home or that the child is not likely to make an improvement in school, they might not be placed in the right situation, etc. It doesn't seem fair to evaluate the impact of the EP in those cases."*(SENCO6)

EPs and SENCOs made reference to the measures that schools were likely to be undertaking with or without the input from the EP. Existing support or the involvements of other professionals were also identified as likely to be important factors in helping to move a child towards a desired target.

> *"It's not always easy to identify what the role of the EP was or is, very difficult, huge overlap between what you [EP] do, what I do and other organisations do."*(SENCO2)

> *"Well, I think it's really hard, isn't it, when you've got so many other agencies working alongside you?"*(EP8)

Interviewees highlighted the subjective nature of progress and, therefore, a strength of TME was perceived to be its applicability to any individual child, which would be reflected in the size and scope of the targets set at baseline (or reset at review). However, owing to the potentially contrasting size and scope of the targets, the difficulty in comparing one TME outcome with another TME outcome for a different child was identified – one may not be comparing like with like (a point that was part of the rationale for developing TME – see Dunsmuir et al, 2009).

> *"I'm too worried about the validity of the whole thing and whether it's able to measure every type of target you might want to measure"*(EP9)

For many interviewees across both groups, using TME to measure the progress of a child in relation to the EP input in the presence of other factors was an *"uncomfortable"* (EP8) step too far. For example, several SENCOs and the majority of EPs reported they felt it was not possible to evaluate outcomes solely on the interventions that were (or sometimes were *not*) put into place, when there may have been many other factors that influenced the child during the intervention period. There was also the identified risk of attributing change solely to the intervention or to the involvement of the EP, when it was perceived there may have been a combination of factors (possibly including that of the EP contribution) that were important, including those beyond the school or educational setting.

> *"Some of it may be initially started by EP but a lot of this is continued through our own systems or from other support services, so there is a lot of cross over."*(SENCO2)

"To be able to evaluate the whole EP service on the basis of those sorts of target would be a really uncomfortable process."(EP8)

The importance of using the same staff member at follow-up as at baseline was a further important consideration, though the interviews suggested this was not always possible owing to the demands on time or availability of the EP. However, interviewees noted the danger of subjective judgements of progress, and the somewhat arbitrary nature of the identified change, where this could only be verified by that one staff member involved with the target setting.

"It's not always easy to communicate with teachers or get feedback, and sometimes it may be difficult for the teacher to be directly involved with the target setting."(SENCO8)

*"Really, you need the TA rather than the SENCO to be doing that. So, the person who is specifically working with the child."(*EP4)

5.3.3 Involving parents and children

A further layer of complexity is added by the potential involvement of parents as well as children in the TME target-setting process. The ease with which parents, in theory, could be assimilated into the TME process was highlighted as a particular strength amongst SENCOs, although it was felt this was not necessarily being put into action.

"TME can help us to find a role for the parents, that they can be involved in working towards the targets and in being an important part of the review process."(SENCO1)

Respondents across both groups highlighted the importance of involving parents in discussions about their children's educational needs. Conversely, however, regularly involving parents in the *TME* planning discussions appeared to be at the early stages, with few references made to parents being present. In addition, several school staff felt that TME on its own might not be enough to satisfy the demands of parents that enough support was being offered to their child.

"There's a bit of pressure to show progress from parents. If we review the TME and progress hasn't been made, well I think they're going to think the whole process was a bit of a waste of time."(SENCO4)

A common theme of interviewee responses in both groups was that of the difficulty in managing parental levels of expectation in support for their child. One SENCO highlighted her frustration at the manner in which school referrals to the EP were sometimes made for children as a result of parental demands. It appears that varying levels of expectation may influence the nature of the EP work and, therefore, the subjective judgement on satisfaction of outcomes using the TME process.

"An EP's intervention is not just a little talk about the TME, it's that the talk that leads to the TME is so much broader, involving the family, expectations, engaging the whole system."(EP10)

"It's even more difficult if it is a parent that's made the request because then nothing might change – it might be more about the parent not accepting the child's difficulty in the first place."(SENCO6)

Involving children in the management and monitoring of targets was identified as a clear theme, with benefits identified for doing so. Responses highlighted the difficulty in using strategies or resources to help a child meet an agreed target, if the child was unaware of what (s)he was aiming to achieve, for example, in terms of motivation or self-esteem.

"There is a by-product of asking for evaluation... which is every time you ask someone for their opinion, you are automatically enhancing their self-esteem, their sense of personal power."(EP10)

"Children had commented that they wanted to be able to say good things about themselves when it came to seeing the EP again."(SENCO3)

While there was an acknowledgement that children *should* be involved in the setting of targets, there was some lack of agreement on how this would be achieved, particularly where the children were perceived as very young or unable to engage in the process. This perceived difficulty may help to illustrate why, despite the broad agreement that it was a necessary and important aim, very few respondents in either sample made reference to children being involved in the monitoring of TME targets. Post-intervention discussion with the children also revealed little awareness of the targets that had been set for them (see Chapter 4).

The *absent* voice of the child thus emerged as an important sub-theme. While reference was made to parents and teachers or support staff being involved in the target setting and reviewing process, not one respondent in either group described children as being present or consulted when targets were set up or reviewed.

> *"If they're not aware of what the target is or they're not involved in the target setting itself you could argue it's a bit of a waste of time."*(EP9)

The implication from the interviewees was that within the TME structure, greater attention needed to be taken to addressing the limitations of current target-setting practice across EPs / schools, and in developing consistent systems to involve children and their families in the process accordingly.

5.4 Superordinate theme 2: Resistance

> *"I'm not sure how much people have been using it. It may have been used because EPs were asked to use it, or they may not have because they were asked to use it."*(EP8)

A number of interview responses from both EPs and SENCOs linked into the subordinate theme of resistance towards TME and in changing working practice. Sub-themes included aspects surrounding motivation towards and engagement with a new system, training and confidence; and the need to establish an appropriate evaluative context within the delivery of EP services. Respondents within the EP group also highlighted aspects of their perceptions of the specific psychological content inherent within the TME framework.

The data are organised in three main categories:

- Perceived value of TME
- Training and confidence
- Establishing an evaluative tool

5.4.1 Perceived value of TME

There seemed to be a discrepancy between the EP and SENCO group responses in their initial perceptions of the TME process. EPs appeared to be more likely to perceive the process negatively. SENCOs meanwhile appeared more likely to perceive TME positively – for example, where it was perceived to promote a sense amongst teachers that change was occurring, even in the most complex, *"stuck"* (SENCO4) cases.

> *"I thought they were very useful, especially for children that seem stuck [and] perhaps maybe don't seem to making progress on NC levels, but actually are."*(SENCO4)

> *"Most [of the schools] have been quite compliant about it, but I get a sense with some it's more they're just pleasing me… They feel it's something I have to have done."*(EP7)

Across both groups there were a number of acknowledgements towards previous EP / school working practices and resistance to change what was perceived as working well. This was particularly so where there was a long established relationship between the EP

and the school, and where TME was regarded as a potential threat to the quantity, quality or type of EP work undertaken.

> *"There may well be a culture of schools not being so clear that target setting and review is part of what they want EPs to do."*(EP8)

EPs and SENCOs raised concern relating to the means by which TME had been imposed as part of routine EP working practice.

> *"It's just another thing to think about, another thing to do."* (EP4)

> *"It feels a bit like an additional layer of paperwork, and not that different to doing an IEP [Individual Education Plan]."* (SENCO6)

Some respondents indicated that they felt TME had been imposed onto the working practice between school and EP without regard to the specific nature of the case work undertaken or referral criteria. A number of SENCOs highlighted the presence of historically changing initiatives in EP service delivery, doubting that TME would be used beyond the short term, and negatively affecting their willingness to learn and become proficient in a new process.

> *"I think that started them off quite badly for me, because the service had been told to use them, there had been no negotiation."*(EP2)

> *"It's like there's too many different agendas within the LA and things change all the time. TME probably won't be here next year."*(SENCO6)

Within the semi-structured interview process it is perhaps not surprising that the EP interviewees had more to say about the evaluation of EP services than did their SENCO counterparts.

> *"It's difficult – it's not an issue I've thought of a lot... I'm not sure if it's something we really do 'officially'."*(SENCO8)

While it was generally acknowledged among the EP group that TME provides some evidence for the success (or not) of EP led interventions, there were some notable exceptions. For example:

> *"You can't evaluate everything, must focus on one area, e.g. "this term we're going to focus on parental feedback and this is what we want to find out. TME is too broad."*(EP6)

While EPs were more likely to address the perceived need to identify a suitable process for evaluation than were their SENCO colleagues, responses amongst the EP group indicated that TME was not wholly regarded as a suitable tool for this purpose. A small number of EPs questioned the psychological value of the TME process, and, therefore, its value and relevance to the nature of the work undertaken.

> *"It's not the sort of thing you can really test. So it's very difficult to measure how well it works."*(EP4)

"There's a high jump between EP work and TME being the EP's outcome. I don't always see the link there, and I'm not sure that they're both measuring the same thing."(EP2)

Concerns were raised with regard to the empirical basis for TME, and to the limited research demonstrating its reliability, validity and suitability for the EP context, although it was also acknowledged that this was a wider issue within the profession as a whole.

"I think we should evaluate, but is this arising out of a defensive modality, which is fine, but would be a bit of a shame if that is what we're driving it on."(EP5)

Other comments alluded to TME being part of a *process* rather than a separate *tool* for evaluating or measuring EP effectiveness. In this regard, the evaluation was described by a small number of EPs as being part of the intervention itself, and, therefore, lacking credibility as a separate evaluation tool. This explanation fits with a number of comments received from EPs who reported that on occasions where TME had been considered for use, the process was not perceived to "fit" the particular piece of (usually individual) casework.

"I've had a case, and quite often it hasn't felt like it suits very well... it's [TME] felt quite complex."(EP2)

"I haven't done a TME because it hasn't been something that [it's] fitted well with... so it would only be evaluating part of that practice."(EP4)

The period of data collection coincided with a range of other changes within the Psychology Service. This included a temporary reduction in the amount of time available to visit schools. It was, therefore, unsurprising that a number of respondents commented on the impact this was having in implementing TME into routine working practice. This included the need to schedule a review despite the pressure in capacity, and perceptions that TME was increasing EP workloads.

> *"Let alone a service saying we haven't got much time so we're going to have to cut it, and at the same time we want to set these targets AND come in a certain time later."*(EP5)

One EP commented that there was maybe a perception amongst both EPs and schools that service priorities were shifting, and that this might explain why more TMEs were not completed within this period.

> *"But the last thing we had was a top priority, and the one before that, so there's that layering of them all being top priorities and not being sure which ones no longer are."*(EP5)

5.4.2 Training and confidence

Respondents amongst both groups reported differing levels of confidence with regard to using TME. While comments were made regarding the apparent simplicity of using TME, there were concerns in both groups around the practical aspects. Several EP respondents were not confident in using TME or were concerned they were not following the same

procedures as other EPs – such as the phase at which the TME form was completed in any given piece of initial casework.

> *"I got a bit confused… maybe I misunderstood?"*(EP7)

In this regard, several EPs commented that they had not received any specific training or supervision in using TME. The SENCOs reported varying degrees of confidence with regard to how well the TME process had been explained to them, with those reporting a higher level of confidence where the rationale and explanation for TME was clearly outlined. Thus the perceived quality of this initial TME interaction between EP and SENCO seemed to be an important consideration for both groups in terms of the perceived likely success of the approach.

> *"I did feel anxious about it… it's been an interesting exercise for somebody who didn't go on the training."*(EP6)

> *"I think it's discussion…the last lot we were all together discussing them and it made a huge difference."*(SENCO4)

Comments indicated that EPs lacked a sense of ownership or empowerment towards TME and, furthermore, the comments from the SENCO's perspective indicated that the value of TME was somewhat inconsistent as presented by different EPs.

> *"We didn't get very much input into it [TME]. So I was left thinking how do I do this?… Whose issue is this?"*(EP5)

"Well we did it. But at first I don't think we were really quite sure why we were doing it."(SENCO3)

The theme was reinforced by those EPs who stated they had not used TME in their casework as much as they should. As described in Chapter 4, the average number of TME forms completed by EPs during the data collection period was 5.7. In many cases, TME was not an approach they had remembered to use, while some admitted they simply did not want to use it. Nevertheless, there was recognition amongst the EP and SENCO samples that it was both desirable and necessary to find a suitable evaluative model to link into a plan-do-review model of service delivery.

"Because it's not easily lending itself, therefore I'm not doing it."(EP4)

"Having a review cycle is quite useful in that sense as it allows a chance to reflect and think again about how to support the children if it wasn't working."(SENCO3)

Some school comments were much more optimistic about the practical application of TME, particularly where this had been used on a more regular basis. For example, in two schools, SENCOs commented that they had requested additional training input from their EP in order to use TME as part of their own in-school monitoring of children with SEN.

"We might try to adapt some of the way we set our IEP targets for all children to use the sort of framework shown by TME, such as using a scale and reviewing it every 6 / 8 weeks."(SENCO5)

"I think teachers also feel quite comfortable about it because it's there or thereabouts in terms of their comfort zone already in terms of target setting."(EP9)

5.4.3 Establishing an appropriate evaluative context

It was acknowledged in both groups that establishing an appropriate evaluative framework was an important principle in EP and school working, specifically for helping to safeguard future EP services through demonstrating value.

"I think it's essential, alongside the measurement of the impact of the other services that come into school, or otherwise it becomes very stagnated if your services don't get evaluated."(SENCO2)

The majority of respondents in both groups were able to identify a range of alternative methods used previously as a means of EP evaluation, or in measuring change or progress towards targets for children. Both EP and SENCO respondents had reservations about TME acting as the sole agent for evaluating the contributions of EPs, with some acknowledgement that while TME could provide some useful information, this should be within the context of a wider evaluative framework.

"It's very difficult to use the same sort of evaluation with every situation... you have to use different approaches depending on what the outcome you're aiming for is".(EP6)

SENCOs identified a range of other evaluation methods that they had liked when it came to thinking about the impact the EP had made. These included, for example,

questionnaires, discussions and informal feedback; and other practical factors – including, it would seem most importantly, the number of EP visits per term (with the implication that *"the more the better"*).

> *"The more EP time you have then it stands to reason that the better the service you are going to receive is and the more positive the outcome for the child and for the school."*(SENCO7)

However, many of the SENCOs were negative about the sorts of evaluation processes they had previously undertaken with regards to the EP service. For example,

> *"I know that I have to do an evaluation of the EP service at the end of the year, but to be honest I don't have a lot of time and often that gets put to the side while I concentrate on the other things."*(SENCO1)

Both groups perceived a weakness in TME providing information only in relation to specific, targeted areas, as agreed at the "baseline" consultation. Both EPs and SENCOs acknowledged that there were likely to be occasions where the targeted TME approach gave rise to benefits not only in the areas being targeted, but also that there were likely to be other beneficial *"knock-ons"* for the child, without the means to measure or capture such data.

> *"So there might be other things… that come out, that aren't actually even being tapped into by TME."*(EP4)

"Also there might be 'knock-ons' which are not being measured by the targets with children, so just by focusing on the intervention in one area, there may be benefits in other areas."(SENCO8)

The responses suggest that perhaps TME does not have to *replace* "traditional" evaluative processes but that it can be complementary in offering a (subjective) means of estimating progress over the short term (such as between consultations), where observable changes in objective and normative scores may not have registered. This would also provide the opportunity to highlight qualitative but still significant issues such as enhanced observable motivation, confidence or persistence (etc).

The short-term nature of the TME 6 – 8 week review period was also commented upon, with the implication of seeking a means to determine if any short-term changes are maintained and increased in the longer term.

"The evaluation rigidly within 6/8 weeks... may not always be appropriate, for example the difference may be made within two weeks and TME will not be picking this up"(EP1)

One implication may be that one of the value of TME is in providing the stimulus for focusing attention upon the setting of specific, short-term targets on the part of those sharing a consultation. This is as opposed to a situation where, by default, the child's progress is evaluated according to the general expectation for his or her class- or age-group which may not be reasonable in the light of significant special needs.

5.5 Superordinate theme 3: The practicalities of TME

"I think it probably is quite difficult to evaluate EP services, but at the same time I think it's probably quite important to try." (EP9)

This theme concerns the use of TME as a tool for evaluating the impact of EP contributions. Perhaps owing to the nature of the interview subject areas and the first-hand experiences of those EPs and SENCOs in using TME, a high number of comments from both groups linked directly into the superordinate theme of using TME as an evaluation tool. Sub-themes within this section were linked to the practical applicability of using TME within the school context, the apparent strengths and weaknesses of using TME as a framework for planning interventions (though not as a specific evaluation tool for EP effectiveness), and divergent views on what TME was measuring and, therefore, for what purposes.

These data are organised in three main sub-themes:

- Applicability in the school context
- A framework for planning interventions
- The purpose of TME

5.5.1 Applicability within the school context

Within both the EP and SENCO responses, many comments were raised with regard to the practical application of TME. A frequently occurring sub-theme related to the ease and simplicity with which the TME form could be completed. Positive comments from SENCOs were offered in relation to the arrangement of the form, with all the information fitting onto

one side of paper. Several SENCOs commented on how this was a possible useful alternative to traditional EP reports, which were described by one SENCO as sometimes being, *"too long and complicated"* (SENCO4), and that the TME form could be easily shared with other teaching or support staff, and parents.

> *"Now the beauty of that is that in a very brief form, I find it very user-friendly."* (EP10)

> *"It's simple enough to easily share with staff and parents."* (SENCO6)

From an EP perspective, while the form was seen as simple, there was an awareness of the *"danger"* (EP7) of *over*-simplicity. Thus, EP comments were at odds with many of the SENCOs in this regard, with the desire for a greater level of information required on the form. Nevertheless, several EPs commented on the practical use of the TME form as a framework to report a school, or home / school consultation meeting.

> *"I think there's a danger in saying TME is all we have, people might be left wondering, 'is that all we have'."* (EP7)

> *"You have both quantitative and qualitative data, clear and easy to read and share."* (EP10)

Within both groups there was a sense that TME was useful as a practical tool to automatically trigger a case review process, perceived as being useful in increasing the likelihood of the agreed interventions taking place. From the EP perspective, some respondents commented favourably on how the TME process fitted into the need to

ensure, for example, regular Common Assessment Framework (CAF) review meetings were being scheduled (although, as reported above, this may be some source of tension within both groups, since other respondents saw the need to schedule a review meeting as time consuming and a potential waste of scarce EP resources).

> *"It supports the notion of having to review, something we need to be encouraged to do more of with schools."*(EP7)

> *"The TME approach is helpful in bringing all the information about a child together into one place, and linking back to the CAF."*(SENCO2)

Time was considered a real challenge to the likelihood of TME taking place successfully or not. SENCOs reported that it was difficult to access EPs, and that scheduling a review meeting between 6 – 8 weeks later was often not practical, particularly if there was a perceived need for the EP time to be prioritised towards other children or situations arising. Both EPs and schools reported that some review meetings had to be postponed owing to illness or other events, and in some cases a review had not been rescheduled by the point at which the interview took place. However, one EP did report that in order to avoid some of the practical difficulties associated with time and scheduling of meetings, she had completed the review over the telephone, and was keen to explore other methods by which IT technology could be used to facilitate the process.

> *"... Like this year when we haven't really always had EP cover, it becomes more difficult when the network breaks down."*(SENCO1)

"Inevitably, if you have implications for time, you have implications for the number of kids you can see."(EP4)

TME was frequently linked into the more established consultation-based means of EP service delivery. Both EPs and SENCOs commented on the clarity of the initial consultation being integral to a successful intervention, with TME a potentially useful framework to achieve this. Although some SENCOs reported the difficulty they had in freeing up teaching or support staff time to meet with the EP, respondents in both groups reported that it was much more useful to have the 'problem-holder' present (usually the class teacher or TA) when setting up the intervention through the TME framework. Comments suggested that this allowed for a richer discussion around the nature of the difficulties and the practical aspects of the intervention to suit the needs of the individual child concerned.

"The TME approach of going through TAs who are directly involved daily with the children has been very positive."(SENCO3)

"That the spark of asking the teacher to work towards specific targets of goals, we are also educating them, we are also advising them, we are used as consultants."(EP10)

Nevertheless, even where all the perceived useful steps were in place, both SENCOs and EPs reported review meetings that had been cancelled because the intervention had not taken place. A combination of reasons was provided, focusing primarily on staff or child sickness, time pressure, or other events taking place (such as school trips etc). One

SENCO admitted that the TME targets had been reviewed even where the agreed intervention had not taken place at all.

> *"We've had lots of absence and sickness, and not being helped by the emergence of swine flu either. This means that putting a programme of support together and sticking to it has been really difficult."*(SENCO3)

> *"But the other issue is getting the review meetings done and effective and with the right people."*(EP4)

Despite the practical difficulties encountered, several of the EP respondents commented that the TME evaluation was useful as a means of examining the circumstances impacting on the situation concerned. The perception was that the review meeting, if it did take place, was then a useful point to reflect on what had happened and to re-think the intervention(s) and the resources available.

> *"Lots of things have happened both in school and at home for the children concerned, TME gives a realistic snapshot of what has happened over that 8 week time period even if it isn't exactly what was planned."*(SENCO3)

> *"It doesn't matter [if the intervention did not take place as planned]... That's what psychologists should be interested in, what are the human factors that facilitate or block change?"*(EP10)

5.5.2 A framework for planning interventions

During the analysis, a number of sub-themes emerged which were more closely aligned with TME as a framework for planning and developing interventions, as opposed to TME being used as a tool to evaluate the work of EPs involved in these interventions. For example, amongst both the sample of EPs and SENCOs, there was an emergent sense that TME was helpful in defining 'the problem', and providing a sense of ownership over the process of defining interventions and strategies with the problem holder (often the class teacher or the TA).

> *"So I think it's possible to map a problem on the TME form, so I can see the format could be useful."*(EP2)

> *"I'm enthusiastic about the by-products of the TME – finding ways to overcome problems and also the motivation for class teachers or TAs."*(EP10)

In turn, this was seen as beneficial in encouraging staff to follow-through whatever the agreed actions were:

> *"[The] reviewing mechanism ensures that there's an expectation that [the intervention] is carried out."*(EP7)

> *"TME data can also be shared across different staff who work with a child to show what is being focused on."*(SENCO8)

SENCOs, in particular, commented on the TME process as being a useful tool, and especially helpful in reducing complexity when addressing difficult situations or planning interventions for children. SENCOs made reference to the *"security"* afforded by TME, and the manner in which it instilled a sense of events *"feeling more manageable now I can see it laid out on the TME form" (comments from focus group).* Both EPs and SENCOs reflected on how the TME review process was useful in exploring what had taken place during the intervention period, particularly in identifying what seemed to be working or where there may be a need for a further, or an entirely different, intervention.

> *"It aids the teacher feeling they now have a better understanding of the needs of that particular pupil."*(EP7)

> *"Often where there are complex needs... it's quite good, it helps to focus on one thing at a time."*(EP4)

Indeed, while the likely complexity of many of the cases was recognised, it appears TME may assist in focusing attention upon prioritising a manageable number of targets (which would not preclude other references to additional benefits), and upon identifying targets that are meaningful and realistic. It may well be logical that this is indeed as much the school's as the EP's responsibility, but what matters is that the process would be part of a *shared* consultation regardless of relative contributions, and that *agreed* targets (and the associated strategies) are more likely to be pursued if part of the TME process involves setting a time scale for the review and an implicit agenda. Therefore, TME appears helpful in directing attention to what happens *between* consultations.

SENCOs commented frequently on how the TME process felt very similar to the IEP process of setting and reviewing targets, and some EPs highlighted the link between the two:

> *"On the whole, I think you'll find it's mostly what we've got on our IEP. I think it's more another version of an IEP."*(SENCO7)

> *"I think for some schools it fits in [with] the IEP system... actually helps to focus more on what they're doing with that student."*(EP7)

However, it was generally agreed that the nature of the targets was usually more specific, and the "expected progress" helped to judge whether or not targets had been met, by helping to create realistic expectations. In one school where provision mapping was used instead of IEPs, the SENCO reported favourably in using TME to link into the resources and support available. Two further SENCOs suggested they would be amending their IEP process to a system more aligned with TME for all the children on the SEN register.

> *"It might be quite helpful in encouraging schools to set more specific and measurable targets than they might previously been doing on an IEP."* (EP4)

It appears that TME does overlap with IEP setting and reviews, but one might usefully specify what TME can add, for example, to the focus upon progress over the short term. This is often qualitative and related to motivational or affective aspects of the child's performance, making it important to capture but not readily measured.

SENCOs and EPs reflected that it was useful to be able to make a judgement using TME about which Every Child Matters (ECM) outcomes were being targeted through the interventions and targets set-up.

> *"It's having to prove that you're doing something that definitely adds value that drives everyone's agenda."*(SENCO6)

However, there were concerns raised within the EP sample about the *"somewhat artificial construction"* (EP5) of the ECM outcomes, with views that different perceptions were likely among different people about what each meant, and the consequent difficulty in making judgements about which category a target best fitted. In this regard, several EPs felt that the interventions set-up were likely to have a more widespread impact on the child than simply in one outcome area, and in these cases all the ECM outcomes had been selected.

Both EPs and SENCOs reported favourably on the use of the 10-point Likert scale, with specific regard to it being clear and useful to document the evidence of what interventions have been used and to what degree of success. Nevertheless, there were difficulties acknowledged with the individual and possibly subjective nature of the target setting process.

> *"Because different people are going to have completely different feelings about what is a successful outcome."*(EP8)

> *"But often progress is just observed isn't it? So it can be quite anecdotal which is hard to measure."*(SENCO2)

EPs and SENCOs highlighted the difficulty using traditional or more objective forms of assessment (such as NC levels or standardised assessment) to measure progress for many of the children for whom TME targets had been set. The individual nature of the targets and criteria for progress was regarded as a useful means of demonstrating and documenting progress, even where it might not be expected that any such progress would be noticeable for the same children using a more general assessment approach.

> *"I do particularly like the scaling, especially for people who can get so entrenched by difficulty."*(EP6)

> *"It does help to show progress, so maybe that's the most important bit."*(SENCO4)

Although all but one of the SENCOs in the study reflected that they were experienced in setting 'SMART' targets for children, it is interesting to compare the relatively poor scores for target quality identified in the quantitative results (Chapter 4) with the comments from SENCOs and EPs in the interviews. While the SENCOs reported that they were generally proficient in setting and reviewing appropriate targets, some acknowledged that it was easier to set and measure certain "types" of target than others. For example, there was a common perception that "learning" based targets were easier to measure than "behaviour" type targets (and that any steps in learning were usually more readily observable)

> *"I'm not sure how easy it would be to do this with some other children in the school where the problems are more around behaviour than learning or communication."*(SENCO6)

"I'm less clear about how we are convincingly using it, I'm just thinking because behaviour is less easily measurable."(EP10)

While the majority of the EPs regarded target setting as an important undertaking, nevertheless, several commented on the difficulty they had setting appropriate 'SMART' targets with the SENCOs or other school staff. The range of comments offered suggested that agreeing and setting out appropriate targets was perhaps one of the most difficult yet crucial parts of the TME process, especially in cases where limited time was available for discussion.

"I think the schools find that quite hard."(EP4)

"I could sit there and write a million targets, all of which are going to be SMART, but it's about the interaction and how you achieve that in a consultation with the stakeholders."(EP9)

5.5.3 The purpose of TME

Descriptions of the utility of TME varied substantially across both groups. However an important distinction was being clear about what information was expected from TME. Broadly, as a tool to measure and evaluate EP interventions, the utility (especially among the EP sample) was low, yet as a process for consultation and planning interventions the utility was regarded as much higher. Therefore, the need to be clear about what was being measured was an important judgement in assessing the purpose and usefulness of TME.

"Once we are clear about exactly what we are measuring and we are honest about what we're measuring, it's a very helpful tool."(EP10)

"So it's really important to be aware of what you're measuring and what you're not measuring."(EP5)

Amongst the sample of SENCOs there was a broader sense that TME was a useful means of evaluating change in the child, and that TME could reasonably be used to identify the range and types of intervention that were having a positive effect on a child's learning or behaviour in school. However, in contrast, amongst the sample of EPs, there was a much greater sense of caution, with frequent references to the *"danger"* of identifying change in the child, when it was more likely that what were really being measured were changes in the perception of the problem stake-holder (usually the teacher or TA).

"There is a lot of value in the TME scale though, especially in relation to the behaviour targets, we're able to say "they're getting there."(SENCO6)

"A good target should either be achieved or not achieved. So for me that doesn't measure the target that measures the perception of the target... which is a different measure, but still a valid one."(EP1)

Respondents from both samples commented on the indirect nature of TME as a tool for measuring EP interventions and, therefore, the EP input. Specifically this appeared to be since it was considered rare for EPs themselves to deliver an intervention, and this appeared to be the source of some tension with respondents in the two groups appearing to approach this from different perspectives. For example, SENCOs were primarily

concerned with EPs assuming responsibility for a successful intervention where this was delivered by a member of school staff. In contrast, EPs in many instances were concerned about interventions not being put into place, or delivered inappropriately despite prior discussion and agreement, and how such "within-school" factors would impact upon their own evaluations.

> *"This might not always be measuring exactly what the EP said or did or recommended, but what happens in the school afterwards."*(SENCO8)

> *"We've got no control really over what's happened in the interim period."*(EP4)

However, a small number of EPs did highlight the responsibility the individual EP had in ensuring that interventions were appropriate and realistic. In this sense there was acknowledgement of the EP's direct input on such "indirect" interventions – for example, in delivering training or support to the consultee responsible for the intervention.

> *"It's not just targets that need to be SMART, but you need to make sure that what you set-up to happen next is also realistic."*(EP5)

> *"TME was helpful in us thinking about our training for next term."*(SENCO2)

Across both groups there was an acknowledgement of the subjective nature of the TME ratings, and how it was consequently difficult to make comparisons between cases within and across schools or different EP 'patches'. There was a unifying link across all respondents that the baseline ratings and reviews are highly subjective and difficult to

quantify meaningfully. Furthermore, some respondents from both groups spoke of the artificiality in reducing children's learning needs to scores on a page.

> *"I suppose it's also a bit sad to think about reducing a child to just a page, but then there's so much paperwork involved in all children that maybe that's not such a bad thing."*(SENCO2)

> *"I think the scale, I think it's quite subjective, isn't it? What a teacher might think is a five, I might think is a four."*(EP8)

Many EPs commented that if TME is not useful for evaluating EP contributions, then measuring the change in *perceptions* is in itself a useful exercise, particularly if it enables a problem-holder to look at the situation from a different or more solution-focused perspective. However, a common thread among EPs was to avoid possible "false-positives", where change was attributed to the child, but where what had actually changed was the consultee's perception.

> *"As long as we're clear about what we're reporting that's legitimate. But I don't think we want to be under any illusion that's the same as saying in cases we've been involved in, the children have made progress".*(EP1)

Respondents in both samples highlighted the associated *"pressure"* to demonstrate positive outcomes following EP interventions. A small number of SENCOs reported that at the planning stage there was an assumption that progress would be made, and that teachers (particularly if they were inexperienced) or support staff may be susceptible to these expectations. Similarly, several EPs commented that the simple presence of an EP

or parent at the review stage would create pressure to indicate progress. One EP commented that the use of the "expected level of progress" at baseline and review automatically created a sense that progress would be made, and to arrive at a different conclusion at review created a tension that the intervention phase was unsuccessful.

> *"... The teacher might not always admit this to you or to me, so the results might get a bit positively skewed sometimes."*(SENCO1)

> *"Teachers know their child has to make progress so they'll say it... like it's almost a leading question."*(EP2)

Furthermore, several EPs commented that there was also a risk of *"false-negatives"*, whereby there might be a *"hidden agenda"* (EP2) or other reason why the school may not wish to acknowledge progress, which may affect the criteria the EP is being evaluated against, and where *"demonstrating progress may be at odds with the original referral criteria"* (EP2).

> *"It might not work because they [schools] want assessments, statutory work, they want all of that, but in an initial TME that might not be discussed."*(EP6)

> *"Certainly in my experience the schools say I like this model and I've seen progress but I need a statutory assessment."*(EP5)

Nevertheless, in the context of anxiety about *perceptions* as much as objectively observable progress, one might suggest that teacher or SENCO anxiety is likely to be a major influence in the setting up of an intervention involving shared consultation. Thus

accordingly, a reduction in anxiety, or a perception on the part of the same teacher or SENCO that progress is being achieved, could be taken as a positive indicator.

Similarly, a perception of positive change in the child is not to be directly equated with a positive evaluation of the input of the EP. However, it may well offer a marker or indication of the EP's contribution given the probability that the EP will have played a leading role in the consultation and planning meeting(s).

5.6 Summary of EP and SENCO comments

EPs and SENCOs had much to contribute to the theme of the role of the EP, evaluation, and TME as a suitable method to achieve this. There were a range of different views, although there was a broad consensus in both groups that it had a considerable degree of utility as a tool for identifying the problem and agreeing interventions. From the SENCO perspective, TME was particularly well regarded for its simplicity, ease of use and clear manner in which it highlighted progress (or at least perceptions of progress).

> *"It provides an indication of progress made even if it's in small steps [and] even if they're small targets. Yep, it seems to have worked well where we've used it."*(SENCO5)

However, this was not regarded as the same process as that required for measuring the specific impact of the EP, and indeed using TME for this purpose seemed to be the source of some tension for both groups in respect of instances where the EP had not been

involved directly with the intervention. There was a sense that TME was not a standalone framework by which EP work could be evaluated.

> *"But in terms of being able to evaluate for every single child the impact of an EP intervention, it's not going to provide the complete picture."*(SENCO4)

The value of EPs working directly through TME with class teachers, especially for TAs and other support staff, was highly regarded, with particular focus on promoting ownership, motivation and helping to increase skills and expertise when working with children and young people.

> *"I think it's useable and useful, there's no reason not to be doing it at all."*(EP5)

There was less consensus among EPs when discussing TME as a suitable means by which to measure the specific psychological contributions they had made. The presence of other factors, and the difficulty in selecting appropriate targets were particularly challenging aspects, alongside the awareness that what was really being measured by TME was changes in stakeholder perceptions, and that this may be highly subjective.

Nevertheless, it appears that the majority of respondents felt TME added some value, and that it should be recognised that TME as a part of regular EP / school interactions was still at a relatively early stage in development.

> *"It might not work for every piece of casework, but I think it's going to have a key role and it probably does have a relevance across our service."*(EP9)

"But then maybe that's because it's early days, I don't know."(EP4)

In this sense, it appears that TME was regarded as adding value simply in terms of it being a better option than not using TME:

"We need to do something, and I would rather be doing something like that [TME] than talking about evaluation and then actually never arriving at what we are going to do in the meantime."(EP5)

The responses received among both groups indicate the perceived complexity surrounding the evaluation of TME. It is interesting that one of the key drivers for this research – the accuracy of measurement through TME – was only a small strand amongst all of this, implying that for the professionals involved, evaluation of EP activity, and even of interventions, ultimately has little to do with demonstrating a child's progress in a reliable fashion.

6.0 Discussion

This chapter provides a summary of the points raised in the qualitative and quantitative analyses presented in Chapters 3, 4 and 5. Each of the research questions is re-visited, in order to explore what conclusions have been drawn from the research, and how these link back to the issues raised in the opening chapters. Consideration is given to the researcher's personal reflections and the implications for further research.

6.1 TME outcomes and the research questions

The following section provides a summary in relation to the original research questions, the rationale for which was provided in Chapter 1.

6.1.1 Research Question 1: How well does TME operate as a means of assessing the objective impact of interventions for an EP Service?

From an evidence-based perspective, it may seem logical that the key element of EP evaluation ought to be based upon successful outcomes for children. In fact, according to the comments received from SENCOs and EPs, there were many bigger questions about such evaluations, for example, the nature of the EP's role, the difficulty in separating elements of influence, and the delivery of services through others. Indeed as discussed earlier, from the comments received, the evaluation of EP activity appeared to have little to do with, ultimately, demonstrating progress on the part of the child in a reliable fashion. Furthermore, from the comments received from the children, it is difficult to ascertain any sense of the impact where the children were not explicitly aware of the TME targets set.

From the quantitative-based data collected, difficulties in evaluating the effectiveness of EP-led interventions appear confounded by the range of targets that were not easily translated into "SMART" targets (even within the specified domains of literacy and behaviour). There were implications surrounding the perceived poor quality of the targets adding to the difficulties in identifying a clear difference between baseline and follow-up. It is possible that, although comments regarding the use of a scale were generally favourable, this may have taken attention away from the need to establish clear criteria for measuring change and what this would look like. One implication is for supporting the request of many respondents for further training in the identification of appropriate and high quality targets.

There also appear to be difficulties even where targets *are* SMART. Such instances arose where it was felt by respondents in both groups that the EP (or school) input had resulted in more *general* implications for the child which were not recorded through TME, or where it was felt the recorded progress had little or nothing to do with the EP. Similarly, the outcomes indicate that evaluating within a short time frame may not always be sufficient to make judgements on an intervention or wider piece of EP work.

The research appears to lend further evidence to the issues raised in Chapter 1 regarding school perceptions of the EP role. From the comments received, it appeared that school staff reflected less on the importance of evaluating the EP role than did the EPs themselves. While this may not be especially surprising, there does appear to be some risk in using TME in an evaluative context where the consultee / consultant (EP) relationship is not clear. Thus, while TME may be useful in identifying change in *individual* children, it is less clear how this could meaningfully be used to evaluate the work of the EPs concerned or of the service as a whole.

The rating exercise involved in TME, especially in respect of behaviour, appeared to help stimulate thoughts about the specific performance of the child across the various elements of the school day as the basis for justifying a particular baseline rating and the further rating after a period of intervention. This is somewhat similar to some of the processes within Solution-Focused Brief Therapy (see Ivesen, 2002) where the need is to be clear about what the desired performance would actually be. Indeed, SENCOs commented that what was important in TME was the perception that change *is* occurring, that ownership of the planning and implementation is still with the teacher and TA (with a greater probability that the actions agreed will be followed through), that it helps to identify what works, and that it ensures a review process.

However, the discussions with TME users revealed some continuing issues such as inter-professional differences about the required quality of targets, the need to involve the parents, children and other stakeholders involved in the intervention (such as class teachers or TAs) more fully in the planning and review processes, and some tensions about the time involved and about the possibly reduced distinctiveness of the various professional roles. Perhaps what matters is time and experience of using this relatively new approach, and allowing the systems involved to develop according to local need and preferences.

6.1.2 Research Question 2: What are the objective outcomes when a more established and objective form of evaluation is used?

In almost all cases, where positive progress was noted using TME, this was also observed using the more conventional forms of evaluation. However, there were inconsistencies in

relation to the level of change in each case – i.e. the degree to which targets had been achieved. However, since TME is reliant on individual, subjective measures of change, perhaps this is not surprising.

The outcomes suggest that in setting up appropriate interventions, whatever process is used to measure progress or change, there will remain other variables that will affect outcomes. Many were highlighted in the research, including personal motivation, training, practice and experience, and the child's own awareness of their targets (and, perhaps, whether or not he or she shared the same enthusiasm for achieving these). However, it seems that TME could have value in highlighting and helping consultees to recognise these factors. Furthermore, using TME as a consultative tool for EPs may be helpful in ensuring that consultees become more focused on actively seeking examples of the target behaviour, and reducing currently observable emphasis on the difficulties arising from the problem behaviour.

The Learning Checklist provided some opportunity to compare progress in the children across the conditions, and offered some, albeit limited, evidence that a positive shift did occur. There were implications, however, for the rate of progress, with children in the behaviour condition likely to achieve below those in the literacy condition on the checklist at baseline, and follow-up. Thus the research suggests the children in the literacy condition may have had a "head start" with regard to likely outcomes compared to those in the behaviour condition, with implications for monitoring whether the right interventions are being put into place for all children. TME appears to be a useful tool for making individual judgements on the nature of interventions within a relatively short time frame for this purpose.

6.1.3 Research Question 3: How does the outcome of TME compare / contrast with the outcome of more established forms of evaluation?

There were clear differences between the levels of progress as measured by TME and the levels according to the more conventional measures. For example, within the behaviour condition, the teachers may have been focusing on different aspects of behaviour than those measured through the external observation. Therefore, it is quite reasonable that the TME outcomes did not show any correlation with the (researcher-led) observation outcomes. This highlights the risk that, no matter how SMART the target, since TME is concerned with individual (usually teacher or possibly TA) *perceptions* of change, EPs and school staff may be focusing on different aspects of behaviour or learning, and thus have different criteria for judging target accomplishment.

In addition, it is difficult to be certain what the criteria are by which school staff measure progress, even where a target has been agreed and shared. For example, in the behaviour condition, there appeared to be a correlation between teachers' perceptions of change and observations indicating that the child was spending more time specifically task-focused. This suggested that being 'task-focused' was likely to lead to perceptions of change, irrespective of what the actual behavioural target was. Therefore, it is necessary to be cautious in the generalisation of TME outcomes.

The research also highlights the potential for situations in which the more conventional form of assessment suggested that change had taken place, but the TME ratings applied by school staff did not. Although this only occurred in relation to one case, it demonstrates the risk that there may be reluctance to acknowledge change if this is not in keeping with the prevailing perception of a difficult pupil, or where there may be a "hidden" agenda

about involving the EP. While this links back to the need to ensure shared agreement from the outset (for which purpose TME may be useful) it also highlights the danger of identifying false negative or false positive outcomes.

The outcomes suggest that TME may be a better indicator of change for those literacy-based targets than those focusing on behaviour. Not only did schools and EPs reflect on the difficulties in measuring behaviour based targets, but it may be that 'learning'-type targets enable a more precise form of measurement via established means, thus focusing the rating applied via TME more coherently. Literacy based targets may allow for a more focused intervention based on a clearer understanding of the nature of the difficulty, unlike where the target is based on behaviour difficulties, where there may be more variables involved and a greater subjectivity in defining the nature of the problem. For example, as Jones (2003) states:

> *"Cross-cultural studies have indicated that teachers' ideal of the well-adjusted pupil is culture specific."*[p151]

However, perceptions of change in the behaviour condition did generally show some progress, which itself be a helpful outcome in assisting school staff to recognise change, no matter how small. Perhaps by sharing tools (such as the observation schedule used in this research) or by discussing different means by which to measure progress, it may be possible to use TME to assist consultees in managing their own quasi-objective measurements in helping to recognise positive change.

Feuer, Holland, Green, Berthenal and Hemphill (1999) state that:

"Policy makers and educators must take responsibility for determining the degree to which they can tolerate imprecision in testing and linking."[p4]

It is clear that the links between the more conventional measurements and TME were not precise. However, the overall trends suggested that positive progress was noted whichever measurement was used. Therefore, in the absence of any alternative method for assessing change in this way, TME appears to be a useful starting point, providing opportunities to consider and discuss wider factors at all stages, if there is enough time and significance devoted to the initial and follow-up sessions.

6.1.4 Research Question 4: To what extent is there a relationship between the perceived utility of TME and the objectively-measured utility of TME?

User expectations appear to influence significantly perceptions about the utility of a TME approach. The outcomes suggest TME was well regarded as a tool for assisting the process of setting up interventions and as a framework for the discussion at review. TME appeared less well regarded as an evaluative tool to measure outcomes for the EP service, and there were implications for increased support and training.

One important factor appears to have been the level of consequences attached to the TME process, which appeared to vary from school to school and EP. From the responses received there appeared to be varying levels of user enthusiasm, thus where TME was used in a low-stakes fashion, there may have been little incentive to focus on the targets. This contrasted with those cases where there appeared to be enhanced expectations upon, and implementation of, TME, and where there may have been a greater instructional focus on achieving the targets. Another factor relates to the lack of awareness of TME by

the children themselves, and the resultant absence of a measure to understand more fully their own perceptions of progress.

Thus, in exploring the reliability and validity of TME, it appears that it may have a relative strength with regard to practicality (defined by Feuer et al, op cit, as, "*Whether the processes necessary to collect the data and conduct the empirical studies are reasonable and manageable*", p.11), but a lower utility with regard to reliability. For example, many of the confounding effects described in the following quote from Feuer et al (op cit) were also manifest within this exploration of TME, with implications across schools, classes, and EPs:

> *"The reliability of the scoring process depends on such factors as the specificity of the scoring rubric, the rater's level of expertise, the quality of the training provided to the raters, and the extent of monitoring of interrater reliability throughout the scoring process."*[p63]

6.2 Summary of research question outcomes

This research suggests that a complicated picture surrounding the use of TME has been uncovered. How much beneficial difference the TME aspect creates is not clear. It appears that it is valuable in focusing attention upon the creation of appropriate targets, what exactly will be involved in moving a child from their initially-rated level of performance to the targeted level of performance, and ensuring a fine monitoring of steps towards that target. What is important is that (SMART) targets *are* established, that strategies are planned and implemented according to the plan, and that monitoring of actual progress is

maintained, even if, in the short term, it might not be possible to register this progress in the scores of standardised assessment instruments.

In the absence of TME, one might wonder if there would be less focus upon targets, or less specificity of the elements which comprise the "expected" target, and some difficulty in recognising the small increments in performance and/or the need for some modification of the components or details of the intervention "package". TME may prove valuable in creating a greater *structure* for the target setting, strategy selection, and ongoing observations of progress while ensuring the significant involvement of teaching staff in the planning and monitoring process.

TME appears to be regarded as a useful tool for assisting the EP, SENCO, teachers, and other professionals, in liaison with the parents and the child him-or herself wherever possible, for example, in measuring perceptions of change and in defining interventions. However, while TME appears to have utility in this regard, there seems no reason why it should be regarded as a *sole* strategy for measuring the effectiveness of EP interventions as opposed to *one* (albeit primary) strategy. As with any tool, its usefulness and the recourse to alternative tools is likely to vary according to the precise circumstances and context of given cases.

6.3 Reflections on the current research

6.3.1 Attributions for successful TME outcomes

Attribution Theory (Weiner, 1985) emphasises that people are strongly motivated by the pleasant outcome of being able to feel good about themselves. In this regard, people's self-perceptions will strongly influence the way in which they will interpret the success or failure of their current efforts. Attributional biases assume people will interpret their environment in such a way so as to maintain a positive self-image. Thus, successes or failures will be attributed to factors that will enhance self-image and feelings of self-efficacy.

The process of TME promotes a sense of the desirability of change – after all, if the consultee (i.e. the teacher) did not feel that the behaviour needed to change it is unlikely that the EP referral would be completed. So since it seems logical to assume that TME targets relate to behaviour(s) that teachers want to modify, attribution theory would provide one explanation for TME outcomes appearing larger than might otherwise be objectively measured or observed, given that outcome is perceived as the direct result of the effort the teacher (or other staff member) has put into developing and delivering the agreed intervention. In other words, there is a direct personal benefit in the student achieving well, since this provides a 'pleasant outcome' of the teacher (for example) being able to feel good about the role they have played in this outcome. Furthermore, within TME this may be equally applicable to any of those individuals involved in target setting, be they the teacher, the child, the EP, or anybody else. For example, in a description of self-reinforcement behaviour, McNamara (1999) states:

"The achievement of self-set criteria or standards can act as a motivator (or at least reinforce existing motivation) to continue."[p73]

The positive feedback provided through TME (i.e. the perception that what the consultee has done has had a direct result, and that the 'intervention' has worked) is likely to increase motivation to engage with both the process of TME and the intervention itself in the future, since according to attribution theory a person's perception for success or failure will determine the amount of effort the person will continue to expend on that activity. If consultees recognise their own role in helping move the situation forward, then attribution theory would suggest a direct link to increasing motivation since they are more likely to engage in the process again. This suggests a need for TME to become embedded in everyday practice in order to achieve the most positive outcomes.

6.3.2 Practical issues relating to successful TME implementation

In the analysis of the data, it appears there are a number of factors to consider when implementing TME that might be helpful in developing effective practice. The first relates to the need to ensure that systems and support are in place to ensure that the TME process is embedded within everyday EP practice. For example, both EPs and school staff reported that access to appropriate training and consistent implementation for all EPs and schools appears important in getting the process started. Access to a clear and consistent system for recording and logging TMEs undertaken may be helpful in this regard.

In the current research, EPs' methods of recording TME data varied, with some produced electronically and others handwritten (to varying degrees of comprehensibility). It may be helpful for EPs to keep a file with completed first round TMEs for easy access and as a

constant reminder to keep checking back for TMEs requiring a review. The need is to ensure that TMEs do not get displaced and that the reviews take place, as other children in schools are prioritised for EP involvement. The data suggests that there needs to be a system in schools *and* in the EP service to ensure this does not happen (and which might be dependent on shared agreement relating to the value of TME by all stakeholders).

While the linking into CAF review processes was considered as a helpful side-effect of TME, it may be useful to find some means by which to link these processes more formally. For example, a system to log TMEs electronically with an automated alert for the review would be logical, particularly if this is linked to both the consultant and the consultee, thus to involve and motivate both parties in the reviewing process (and to overcome the potential problem of EPs not being called back by schools to complete the review).

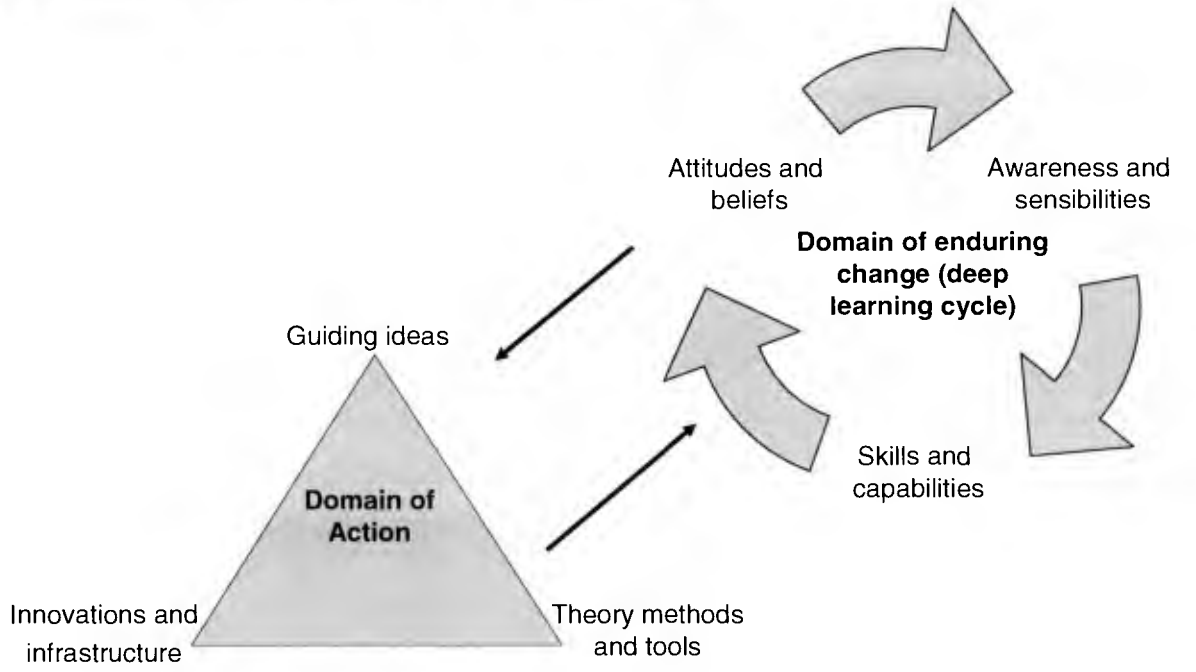
In their text on learning within organisations, Senge et al (1999) describe the need for a process of constant testing of experiences and the transformation of experiences into knowledge accessible to all in order to overcome the barriers to change which may otherwise be consistently reinforced because they are never challenged.

Senge et al (op cit) describe the need to develop "guiding ideas" in relation to stakeholder confusion and questions, such as, "Why are we doing this?" or "What's this change in infrastructure all about?" They describe how,

> *"Time and resources are poured into achieving intended changes. But after a year, with little tangible to show for the effort, something else hot comes along and the effort is abandoned."*[p36]

Senge et al (op cit) propose a model (see **Figure 6.1** below) whereby the guiding ideas; theory, methods and tools; and changes in the infrastructure are all important factors in the implementation of any new system or process, such as EPs using TME consistently in casework with schools. Nevertheless, they propose that it is also the more subtle (and possibly hidden) *"deep learning cycle"* that is equally important in maintaining change, consisting of individual attitudes and beliefs; skills and capabilities, and awareness and sensibilities.

Figure 6.1 Senge et al (1999)'s Deep Learning Cycle



Within any system, Senge et al (op cit) highlight the need to think both about those covered within the 'triangle', as well as those within the 'circle'. In applying this principle to TME, having the TME framework, policy and a member of staff responsible for the implementation may be important. However, there is also the need to reflect on other aspects of practice to embed TME within routine working, through, for example, opportunities for further discussion and innovations, involving plans, resources, training and shared ideas.

Senge et al (op cit) also highlight how the results of change should not be measured prematurely, since this may lead to erroneous conclusions about success. However, and with particular regard to the Local Authority Children's Service context, where there is a very real need to demonstrate evidence quickly, Senge et al (op cit) state that:

> *"This principle, while easy to state, can be very difficult for impatient managers and organizations to practise."*[p45]

Nevertheless, as has been previously highlighted, there is a need to allow time for change to become embedded within practice, with the implication, perhaps that the use of TME should continue to be monitored carefully. For example, O'Brien (in Senge et al, op cit) states that:

> *"Time periods for measurement must be congruent with the gestation period of the learning."*[p45]

6.4 Implications for the current researcher

Taylor and Ussher (2001) describe the active role the researcher takes when identifying themes or outcomes, and in choosing which to report. In my analysis, I was aware of my active role at each step in the research process, including design, implementation and interpretation level.

Elliott, Fischer and Rennie (1999) highlight the importance of allowing the reader to consider issues of researcher experience and perspectives when presenting research findings, thus allowing the reader to consider for themselves how this may have impacted on the final analysis of data. My experience of working within the EP service had provided me with an insight into the topic and associated difficulties of finding an appropriate method for systematic EP casework evaluation. I was also aware that not only had the subject of EP evaluation been raised as a priority concern within my own employing EP service, but also at national level, and across other services within the multi-agency CS context. It was the recognition of the complexity of these issues that helped motivate me to undertake this research, in order to further my understanding of such a complex demand and to add to the body of research within this field.

During all phases of the research, I was aware of my role within the EP service where the research was taking place. I was aware that my status as a team member may have impacted on what the EP interview participants were willing to contribute, and how my own experiences may have affected my subsequent analysis. Furthermore, I also had experience of using TME in my own case work. This could be viewed as a source of bias, but I believe also served to increase my sensitivity to the data as a researcher (Strauss and Corbin, 1998).

These experiences and motivations were carefully considered at all stages of this research. I made sure that, as suggested by Braun and Clarke (2006) I kept comprehensive, reflective notes at each juncture, in order to monitor the impact of my experiences. These notes were frequently used as the starting point for discussion with my research supervisors.

6.4.1 Additional impacts on the research

During the period of data collection, there were a number of additional factors which may have impacted on EPs' priorities, and, therefore, willingness and capacity to implement TME. For example, owing to the financial situation affecting many LAs at this time, the service was going through a process of considerable re-structuring. At the same time, a number of other new initiatives were being introduced including a new electronic CAF system, which had an impact on the number of new referrals coming into the service at this time (and, therefore, those cases for inclusion in the current research).

It is possible that the changes within the EP service had an influence on the interviews, particularly those completed in schools (where EP time had been cut by up to half). The interviews may, therefore, have been the first opportunity for staff to 'have their say' on these developments.

With regard to cases for inclusion in the research, there were difficulties related to EPs not being able to review cases within the time scales. While some of these related to child illness or absence from school, there were other problems, such as unexpected events (for example, critical incidents) impacting on EPs not being able to make the necessary arrangements within time frames or before the six week summer holiday. Further, some schools were unwilling to host a review meeting where they felt the interventions (for various reasons – but usually involving staff or child illness) had not been implemented and wanted additional time to, in effect, begin the process afresh.

In the research process itself, there were a total of five additional cases where it was not possible to follow-up the baseline assessment or observation. These cases were excluded

from the analysis. This included two instances where the children left school, two owing to prolonged absence and one owing to the child having been excluded (and where it might be hypothesised that the TME outcome, were it to have been reviewed, would indicate worse than expected progress). While the intention was to complete follow-up observations at the same time and in the same timetabled lesson as had been the case at baseline, this was not always possible owing, for example, to illness or school trips. Therefore, on these cases the review was held on the next available slot, usually within one day of the target.

6.4.2 Impacts on interviews with children

Interviews were not completed with *all* children in the study for a combination of factors. In some cases this was time related where, for example, the observation or assessment led directly into playtime or lunch, or the children were taking part in another activity. In two cases it was judged by the teacher that the children would not want to engage with me separately from the class, and it was not appropriate to ask the questions in the presence of the other children. Furthermore, as Robson (2002) states:

> *"Interactions between interviewer and interviewee can also be influential."*[p252]

Therefore, even where they did take place, children may have been reluctant to engage with follow-up interviews owing to the unfamiliarity of the situation and researcher.

6.4.3 Data collection method

As described earlier, it is likely that in the behaviour condition, consultees may have been focusing on different elements of the child's behaviour than those of the schedule itself. In the literacy condition, there was also likely to be differences between what the targets were aiming to achieve and what could be accessed via the YARC literacy-based assessment. For example, the YARC coverage was over a set of sub-domains looking at particular elements of literacy. However, this focus may not have been specific enough. Nevertheless, it is unlikely that any single and effective tool could be identified, owing to the individual nature of TME, as argued in this quote from Feuer (op cit):

> *"No test can possibly tap all the concepts and processes embedded in a subject area as vast as reading"*[p67].

Furthermore, there are likely to be difficulties with the measurement of skills regardless of how this is assessed, be this by test or assessment. Feuer et al (op cit) state:

> *"If a test taker also took an equivalent, but not identical, test on a different day in a different place, her score is unlikely to be the same".*[p88]

Not only do the tests themselves have a margin for error, but so, presumably, does the person assigning a TME rating. A further research interest might involve investigating whether outcome ratings would be consistent over time (whereas the present research found evidence that perceptions of targets varied from person to person, and across schools and EPs).

6.4.4 Implications for further research

Feuer et al (op cit) state the need for extensive research when considering any new form of assessment or testing in order to:

> *"Determine the level of precision needed to make valid inferences between tests... [to be sure about] how important are the differences? Can these be overlooked?"*[p93]

Further research investigating TME may helpfully look in more detail at the precise calibration between some form of more objective or standardised assessments and the TME ratings themselves, perhaps by focusing much more specifically on one sub-domain. This would help provide further evidence relating to the appropriateness of assigning a quantitative outcome to the largely qualitative process of TME, and to investigate its validity further.

While TME may have utility as a process for setting up interventions and managing individual change, the danger of overly relying on quantitative measurement and ignoring the individual qualitative elements is emphasised by O'Brien (op cit):

> *"[Organizational] cultures that are saturated exclusively in scientific principles have an insatiable appetite for quantitative measurement – even where they misrepresent truth and reality… There are times when the organization would have been better off without a measurement than with a faulty one".*[p46]

In the current study, there were difficulties reported with regard to time, and the need to hold TME consultation meetings face-to-face. It would be interesting to compare the

outcomes where alternative forms of communication are used, such as the telephone or email, and particularly with regard to follow-up discussions. This may help to relieve some of the pressures on both EP and school time. It would also be useful to compare outcomes where parents and / or children were involved consistently in the setting up and reviewing of targets.

Finally, there were difficulties associated with the setting up of TMEs towards the end of term and in proximity to a school holiday. Consequently, the majority of TMEs were set up in the first week or two of a half term, and reviewed towards the end. In some cases the process crossed over the one week half term break in the summer term, but no TMEs were submitted for review if this breached either the spring or summer extended holidays. Further research would be useful in examining the impact of such breaks from school in terms of the impact on the intervention between baseline and follow-up, but also to examine for any negative (i.e. loss of skill) effects on previously accomplished TME targets, in order to help identify what factors are important in maintaining TME outcomes.

7.0 Conclusions

Anderson, Blumenfeld, Pintrich, Clark, Marx and Peterson (1995) discuss the advantages in developing a model for bringing psychology into the school setting:

> *"This advantage is not only afforded by mere knowledge about concepts, principles and theories; it is only manifested when those ideas are tied together as coherent frames that suggest when and how the ideas can be use."*[p145]

Throughout the interviews, reference was made to the appropriateness of TME as a tool or framework for planning and managing interventions in schools. The interview responses suggest that it has positive utility, especially where linked to the existent consultative model of EP service delivery (see, for example, Wagner, 1995). Fredericksen and Cameron (1998) highlight the difficulties in which:

> *"A teacher may have come to accept that there is nothing that a school can do to help a pupil with specific learning difficulties, or a parent may have become resigned to accepting a child's aggressive behaviour".*[p10]

TME may, therefore, be helpful in illustrating change and moving forward 'stuck' cases, and where progress according to more conventional or National Curriculum-based assessments may be negligible or difficult to identify through conventional methods. TME may also be regarded as tool for EPs and teachers, with implications for motivation and ownership of problems, and to enable them to:

"Become convinced of their powers as agents of change, shed elements of unhelpful practice, develop and research innovative approaches to problem management and demonstrate the effectiveness of what they are doing to increase pupil attainment." [Fredericksen and Cameron, 1998, p11]

In seeking to evaluate the impact of the EP, the research highlights the need to be clear from the outset exactly what the EP role entails – for example, through determining what will be the precise nature of the relationship between school and EP. Magi and Kikas (2009) describe how there has been a lack of change in the roles and functions of EPs in relation to the nature of expectations maintained by school staff. Linked to the findings of Ashton (2006) described in Chapter 1, Magi and Kikas (op cit) found that while various surveys can be cited to indicate that teachers generally value the work of EPs, and desire further EP time for individual consultation and preventative work, this should not be at the expense of assessing children who may have SEN. School staff identified TME as a potential threat to the maintenance of this form of EP working.

Magi and Kikas (op cit) conclude that one of the major inhibitory factors preventing EPs from developing casework approaches in schools is the lack of stakeholder knowledge about the potential benefits of alternative ways of working. Nevertheless, from the responses received from school staff, it seems that in particular regard to funding (such as having schools "buy-back" EP time) it is necessary to be sensitive to the expectations of clients. Perhaps what matters is that the work of the EP complements and extends the existing level of working and that there is genuine partnership between school staff and EPs. TME appears to be a useful tool in helping to generate such shared understanding within individual cases, and as a process for highlighting the individual change therein.

TME was generally well regarded by interviewees in schools, particularly with regard to its simplicity, ease of use and clear manner in which it can be used to evidence progress. It seems the direct work with class teachers, and especially for TAs and other support staff, is valued - promoting ownership, motivation and an increase in skills. While it could be hypothesised that involving children in the target-setting process may have a comparable impact on motivation and ownership, the current research was not able to investigate this further since the interviews with children indicated they had not been routinely consulted as part of the TME process. Involving children in this way emerges as a clear recommendation for future implementation of TME from this research.

TME appears to fit well within a plan-do-review model for EP based casework, and it seems to provide some useful information with regard to *individual* change. Perhaps it is not surprising that TME was seen by some EPs as 'better than doing nothing' with regard to evaluation of services as a whole. Nevertheless Feuer et al (op cit) state:

> *"Error in the linkage between tests and assessments can remain hidden from immediate view unless serious efforts are made to ferret them out."*[p47]

Such a view emphasises the risk also evidenced within this research that at surface level conclusions or associations may be made (or not made) between *actual* performance and *perceived* performance (as measured through TME), which may or may not be real. This necessitates the need for on-going and routine monitoring, assessment and checking back (etc.) with the pupils concerned, as part of a plan-do-review model of service delivery, and to embed TME firmly as *one* means for gathering whole- service evaluative data but not the *unique* means.

7.1 Final summary

This study has involved a small scale evaluation of the use of TME among a sample of EPs, teachers and SENCOs, across a number of schools. In the research, TME has been regarded as valuable by school staff in respect of, for example, providing a structure for EP / school consultations (including establishing roles and expectations), highlighting shared ownership of the issues, assisting in drawing attention to the need to establish realistic targets and upon indications of progress *towards* the targets. Some reservations were expressed by school staff in respect of resistance to established working practice, access to the EP and in identifying the variable (or person) that made the difference.

EP responses favourable towards TME included using it to promote the establishment of school / EP roles and expectations, developing motivation in the consultees to complete interventions, and in using TME as part of a service delivery method involving 'plan-do-review' frameworks. Some concerns were expressed in terms of difficulties in integrating TME with all forms of casework undertaken, the subjective nature of the ratings applied, and the indirect means by which it provided an indication of EP effectiveness (and, therefore, affecting it's suitability as a tool for evaluating the EP service as a whole).

TME and 'traditional' or 'conventional' assessments could be seen as complementary, with TME offering a means of identifying initial progress over a short time scale through observations which are largely qualitative but systematic. Indeed, TME appears helpful as a means of structuring consultations and reviews, and focusing attention upon the initial targets and steps achieved (for which standardised tests may not be appropriate or sufficiently sensitive).

In conclusion, as a structure introduced by the EP service, TME would appear to contribute to determining the significance of EPs' work in respect of encouraging a consensus about the nature of issues, the associated learning or behavioural targets, and of increasing sensitivity to the steps that are being achieved and, accordingly, of enhancement of feelings of efficacy among teaching staff and/or other professionals and parents. The significance of the EPs' particular input may still be judged largely on the basis of outcomes mediated by other people but TME appears to have the power to highlight change and progress, including in attitudes.

References

Association of Educational Psychologists (2005). *NEC Agenda Paper 12.05/16b*

Association of Educational Psychologists (AEP). (2009). *The Evaluation of Educational Psychology Services in the Light of Outcomes for Children*. Association of Educational Psychologists

Anderson, L., Blumenfeld, P., Pintrich, P., Clark, C., Marx, R. and Peterson, P. (1995). Educational psychology for teachers: reforming our courses, rethinking our roles. *Educational Psychologist*, 30(3), 143 – 157

Ashton, R. and Roberts, E. (2006). What is valuable and uniue about the educational psychologist? *Educational Psychology in Practice*, 22(2), 111 – 223

Banister, P., Burman, E., Parker, I., Taylor, M. and Tindall, C. (1994). *Qualitative Methods in Psychology. A Research Guide*. Open University Press

Baxter, J. and Fredericksen, N. (2005). Every Child Matters: Can educational psychology contribute to radical reform? *Educational Psychology in Practice*, 21(2), 87 – 102

Boothroyd, R. and Banks, S. (2006). Assessing outcomes in individually-tailored interventions. *The Lancet*, 367, 801 – 902

Boyatzis, R. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development*. Sage.

Brann, P., Coleman, G. and Luk, E. (2001). Routine outcome measurement in a child and adolescent mental health service: an evaluation of HoNOSCA. *Australian and New Zealand Journal of Psychiatry*, 35, 370 – 376

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77 – 101

Buzan, T. (2000). *Use Your Head.* Ashford Colour Press, UK

Cameron, R. (2006). Educational Psychology: the distinctive contribution. *Educational Psychology in Practice*, 22(4), 289 – 304

Cardillo, J. and Smith, A. (1994). Reliability of Goal Attainment Scores. In Kiresuk, T., Smith, A. and Cardillo, J. (1994). *Goal Attainment Scaling: Applications, Theory, and Measurement.* Lawrence Erlbaum Associates, Publishers

Children Act (2004). http://www.opsi.gov.uk/acts/acts2004/pdf/ukpga_20040031_en.pdf (accessed 12th May 2010)

Children's Workforce Development Council (2010). *http://www.cwdcouncil.org.uk/educational-psychology/*what (accessed 12th May 2010)

Cline, T. (1994). Quality assurance in an Educational Psychology Service, what can we learn from earlier experience in service evaluation? *School Psychology International*, 15(3), 209 – 234

Cormack D. F. (1991). *The Research Process in Nursing 2nd edition.* Blackwell Publishing

Cottrell, D. (2002). Body of evidence: towards evidence based practice. *Young Minds*, 58, 34 – 37

Cytrynbaum, S., Ginath, Y., Birdwell, J. and Brandt, L. (1979). Goal attainment scaling: a critical review. *Evaluation Quarterly*, 3, 5 – 40

Denzin, N. (1978). *The Research Act: A Theoretical Introduction to Sociological Methods.* New York: Praeger

Department for Children, Schools and Families (2001). *Special Educational Needs Code of Practice.* London: DCSF

Department for Education and Employment (DfEE). (2000). *Educational Psychology Services (England): Current Role, Good Practice and Future Directions. The Report of the Working Group.* London: HMSO.

Department for Education and Skills (DfES). (2003). *Every Child Matters.* London: DfES

Dowling, J. and Leibowitz, D. (1994). Evaluation of Educational Psychology Services: past and present. *Educational Psychology in Practice*, 9(4), 241 – 250

Doyle, L., Brady, A. and Byrne, G. (2009). An overview of mixed methods research. *Journal of Research in Nursing*, 14(2), 175 – 185

Dunsmuir, S, Brown, E., Iyadurai, S and Monsen J. (2009). Evidence-based practice and evaluation: from insight to impact. *Educational Psychology in Practice*, 25(1), 53 – 70

Elliott, R., Fischer, C. And Rennie, D. (1999). Evolving guidelines for publication of qualitative research studies in psychology and related fields. *British Journal of Clinical Psychology*, 38, 215 – 229

Evans, H.M. (1981). Psychiatric patient participation in goal setting as related to goal attainment. *Dissertation Abstracts International*, 41(12-B, Pt 1), 4658 – 4659

Farrell, P., Woods, K., Lewis, S., Rooney, S., Squires, G. and O'Connor, M. (2006). *A Review of the Functions and Contributions of Educational Psychologists in England and Wales in the light of "Every Child Matters: Change for Children"*. Nottingham: DfES Publications

Feuer, M., Holland, P., Green, B., Bertenthal, M. and Hemphill, F. (1999). *Uncommon Measures. Equivalence and Linkage Among Educational Tests*. National Academy Press, Washington D.C.

Fielding, M. (1999). Target setting, policy pathology and student perspectives: learning to labour in new times. *Cambridge Journal of Education*, 29(2), 277 – 287

Fox, M. (2003). Opening Pandora's Box: evidence-based practice for educational psychologists. *Educational Psychology in Practice*, 19(2), 91 – 102

Frederickson, N. (2002). Evidence-based practice and educational psychology. *Educational and Child Psychology*, 19(3), 96 – 111

Frederickson, N. and Cameron, S. (1999). *Psychology in Education Portfolio*. Windsor: NFER-Nelson

Friedman, M. (2005). Good is Not Enough. *How to Produce Measurable Improvements for Customers and Communities*. Trafford Publishing: Canada

Gersch, I. (2004). Educational Psychology in an age of uncertainty. *The Psychologist*, 17(3), 142 – 145

Good, R.H., Simmons, D.C. and Smith, S.B. (1998). Effective academic interventions in the United States: evaluating and enhancing the acquisition of early reading skills. *Educational and Child Psychology*, 15(1), 56 – 70

Greene, J., Caracelli, V. and Graham, W. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis* , 11, 255-274

Grey, P. (2001). *How Good is Your Service? Regional Guidance on Evaluating Support Services for Children with Special Educational Needs*. East Midlands SEN Partnership

Gutkin, T. and Curtis, M. (1999). School based consultation theory and practice: the art and science of individual service delivery. In T. Gutkin and C. Reynolds (Eds), *The Handbook of School Psychology Third Edition*. John Wiley and Sons, Inc

Hurn, J., Kneebone, I. And Cropley, M. (2006). Goal setting as an outcome measure: a systematic review. Clinical Rehabilitation, 20, 756 – 772

Ivesen, C. (2002). Solution-focused brief therapy. *Advances in Psychiatric Treatment*, 8(2), 149 – 157

Johnson, R.B. and Christensen, I.B. (2004). *Educational Research: Quantitative, Qualitiative and Mixed Approaches*. Boston, MA: Allyn and Bacon

Johnson, R.B. and Onwuegbuzie, A.J. (2004). Mixed methods research: a research paradigm whose time has come. *Educational Researcher*, 33(7), 14 – 26

Johnson, R.B., Onwuegbuzie, A.J. and Turner, L.A. (2007). Toward a definition of Mixed Methods research. *Journal of Mixed methods Research*, 1(2), 112 – 133

Jones, R. (2003). The construction of emotional and behavioural difficulties. *Educational Psychology in Practice*, 19(2), 147 – 157

Kiresuk, T.J. and Sherman, R.E.(1968). Goal Attainment Scaling: A general method for evaluating community mental health programs. *Community Mental Health Journal*, 4, 443 – 453

Landis, R. and Koch, G. (1977). The measurement of observer agreement for categorical data. *Biometrics*. 33, 159 – 174

Love, P. (2009). 'Educational psychologists: the early search for an identity', *Educational Psychology in Practice*, 25(1), 3 — 8

Magi, K. and Kikas, E. (2009). School psychologists' role in school: expectations of school principals on the work of school psychologists. School Psychology International, Vol. 30, No. 4, 331-346

Marsh, R. (2005). Evidence-based practice for education? *Educational Psychology*, 25(6), 701 – 704

McNamara, E. (1999). *Positive Pupil Management and Motivation.* David Fulton Publishers, London

Monsen, J. and Cameron, R. (2005). Quality psychological advice for teachers, parents / carers and LEA decision-makers with respect to children and young people with Special Educational Needs. *Educational Psychology in Practice*, 21(4), 283 – 306

Morgan, D.L. (2007). Paradigms lost and pragmatism regained: methodological implications of combining qualitative and quantitative methods. *Journal of Mixed Methods Research*, 1, 147 – 163

Newman, T. (1999). Evidence-based child care practice. *Highlight 170.* London: National Children's Bureau

Office of Public Sector Information (2004). *Children Act.* London: OPSI

O'Keefe, D.J. and Medway, F.J. (1997). The application of persuasion research to consultation in school psychology. *Journal of School Psychology*, 35 (2), 173 – 193

Roach, A. and Elliott, S. (2005). Goal attainment scaling: an efficient and effective approach to monitoring student progress. *Teaching Exceptional Children*, 37(4), 8 – 17

Robson, P. (2002). *Real World Research, Second Edition*. Blackwell Publishing

Senge, P., Kleiner, A., Roberts, C., Ross, R., Roth, G. and Smith, B. (1999). *The Fifth Discipline, Fieldbook Strategies and Tools for Building a Learning Organization*. Nicholas Brearley Publishing, London

Smith, A. (1994). Introduction and Overview. In Kiresuk, T., Smith, A. and Cardillo, J. (1994). *Goal Attainment Scaling: Applications, Theory, and Measurement*. Lawrence Erlbaum Associates, Publishers

Snowling, M.J., Stothard, S.E., Clarke, P., Bowyer-Crane, C., Harrington, A., Truelove, E., Nation, K. And Hulme, C. (2009). *York Assessment of Reading Comprehension*. University of York, Centre for Reading and Language

Stoker, R. (2002). Educational Psychology: diversity led, context sensitive, and process driven – a view of the future. *Debate*, 101, 9 – 14

Strauss, A. and Corbin, J. (1998). *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory*. Sage

Taylor, G. and Ussher, J. (2001), Making sense of S&M: a discourse analytic account. *Sexualities*, 4, 293 – 314

Torgerson, C.J. and Torgerson, D.J. (2001). The need for randomised controlled trials in educational research. *British Journal of Educational Studies*, 49(3), 316 – 328

Wagner, P. (1995). *School Consultation: frameworks for the practising Educational Psychologist*. London: Kensington and Chelsea EPS

Wagner, P. (2000). Consultation: developing a comprehensive approach to service delivery. Educational Psychology in Practice, 16(1), 9 – 18

Weiner, B. (1985). An attributional theory of achievement motivation and emotion. *Psychological Review*. 92(4), 548-573

Willig, C. (2001). *Introducing Qualitative Research in Psychology: Adventures in Theory and Method*. Open University Press

Wood, A. (1998). Okay, then, what do EPs do? *Special Children*, May, 11 – 13

Appendix 1: TME form

Pupil:	School:
Consultee:	Date of consultation:
Educational Psychologist:	Date of review:

Target 1: _____

Rating: 1 2 3 4 5 6 7 8 9 10
Descriptor of baseline level:

Descriptor of level achieved:

ECM Outcome: Be Healthy [ ] Stay Safe [ ]
Enjoy & Achieve [ ] Economic Wellbeing [ ]
Make a Positive Contribution [ ]

Target 2: _____

Rating: 1 2 3 4 5 6 7 8 9 10
Descriptor of baseline level:

Descriptor of level achieved:

ECM Outcome: Be Healthy [ ] Stay Safe [ ]
Enjoy & Achieve [ ] Economic Wellbeing [ ]
Make a Positive Contribution [ ]

Target 3: _____

Rating: 1 2 3 4 5 6 7 8 9 10
Descriptor of baseline level:

Descriptor of level achieved:

ECM Outcome: Be Healthy [ ] Stay Safe [ ]
Enjoy & Achieve [ ] Economic Wellbeing [ ]
Make a Positive Contribution [ ]

Appendix 2: Observation schedule

Pupil's Name:	D.O.B:		Year:	
School:	EP:			
Date:	Setting:			

	Baseline / Follow-up (del as appropriate) Observations													
	AET	PET	TDI	OFT V	OFT M	OFT P	TLK	INT	SEP	SEG	RUN	PAG	VTH	OTH
• 1														
• 2														
• 3														
• 4														
• 5														
• 6														
• 7														
• 8														
• 9														
• 10														
• 11														
• 12														
• 13														
• 14														
• 15														
• 16														
• 17														
• 18														
• 19														
• 20														
• 21														
• 22														
• 23														
• 24														
• 25														
• 26														
• 27														
• 28														
• 29														
• 30														

Target Behaviour Frequency Record:	
• Target 1	
• Target 2	
• Target 3	

Other observations:

Key to Observation Schedule

AET Actively Engaged in task
PET Passively Engaged in task
TDI Teacher Directed Instruction
OFT V Off Task – Verbal
OFT M Off Task – Motor
OFT P Off Task – Passive
TLK Talking or Yelling
INT Interrupting Others
SEP Fidgeting in Seat
SEG Getting out of Seat
RUN Running around the classroom
PAG Physical Aggression
VTH Verbal Threats
OTH Other

Appendix 3: Inter-rater analysis

a) Overall observation condition

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Row * Column	85	100.0%	0	.0%	85	100.0%

Row * Column Crosstabulation

Count

		Column		
		1	2	Total
Row	1	71	4	75
	2	1	9	10
	Total	72	13	85

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.749	.106	6.987	.000
	N of Valid Cases	85			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

$$P_o = \frac{\text{Agreements}}{\text{Agreements} + \text{Disagreements}}$$

$$P_o = \frac{71 + 9}{85} = \mathbf{.94}$$

b) On-task behaviour condition

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Row * Column	85	100.0%	0	.0%	85	100.0%

Row * Column Crosstabulation

Count				
		Column		
		1	2	Total
Row	1	44	5	49
	2	6	30	36
	Total	50	35	85

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.734	.075	6.769	.000
	N of Valid Cases	85			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

$$P_o = \frac{\text{Agreements}}{\text{Agreements + Disagreements}}$$

$$P_o = \frac{44+30}{85} = \mathbf{.87}$$

Appendix 4: EP / SENCO Interview schedule

Semi Structured Interview Questions (EPs)

1. **What do you think the purpose of "evaluation" is?**
    a. Broad perspective / first thoughts for follow-up
    b. What is understood by the concept of evaluation?
    c. How far is evaluation an achievable goal for EPs (etc)?
    d. Understanding the purpose of evaluation

2. **Can you tell me about some ways that you feel your work as an Educational Psychologist may appropriately be evaluated?**
    a. Neutral - not focusing on TME yet
    b. Alternative suggestions
    c. Issus / alternatives

3. **Can you tell me about your experiences of using TME within your schools (or other settings)?**
    a. First thoughts on TME

4. **How well do you think TME addresses the need for a framework to evaluate EP interventions?**
    a. Goodness of fit?
    b. Does it address the need?
    c. Specific contexts?
    d. Service wide... local needs? National context?

5. **What processes are required in order to ensure that appropriate targets are set?**
    a. Including TME and targets in general (e.g. IEP)?

6. **How do you know that a target has been achieved?**
    a. Including progression along the Likert scale

7. **Can you describe to me the strengths of adopting a TME-based approach?**
    a. What is it useful for?
    b. Contexts?
    c. Types of intervention?
    d. Complexity of cases?

8. **What are the weaknesses of adopting a TME-based approach?**

9. **Is there anything else you would like to tell me either about the TME process or about the wider theme of evaluating EP interventions?**

10. **In sum, what do you think the utility of TME is, in light of everything else you've said?**

Semi Structured Interview Questions (SENCOs)

1. **What do you think the purpose of "evaluation" is?**
    e. Broad perspective / first thoughts for follow-up
    f. What is understood by the concept of evaluation?
    g. How far is evaluation an achievable goal for EPs (etc)?
    h. Understanding the purpose of evaluation

[Introduce concept of evaluating the work of EPs in schools if not already highlighted]

2. **Can you tell me about some ways that you think the work of Educational Psychologists may appropriately be evaluated?**
    d. Neutral - not focusing on TME yet
    e. Alternative suggestions
    f. Issus / alternatives

3. **Can you tell me about your experiences of using TME within your school?**
    b. First thoughts on TME

4. **How well do you think TME addresses the need for a framework to evaluate EP interventions?**
    e. Goodness of fit?
    f. Does it address the need?
    g. Specific contexts?
    h. Service wide... local needs? National context?

5. **What processes are required in order to ensure that appropriate targets are set?**
    b. Including TME and targets in general (e.g. IEP)?

6. **How do you know that a target has been achieved?**
    b. Including progression along the Likert scale

7. **Can you describe to me the strengths of adopting a TME-based approach?**
    e. What is it useful for?
    f. Contexts?
    g. Types of intervention?
    h. Complexity of cases?

8. **What are the weaknesses of adopting a TME-based approach?**

9. **Is there anything else you would like to tell me either about the TME process or about the wider theme of evaluating EP interventions?**

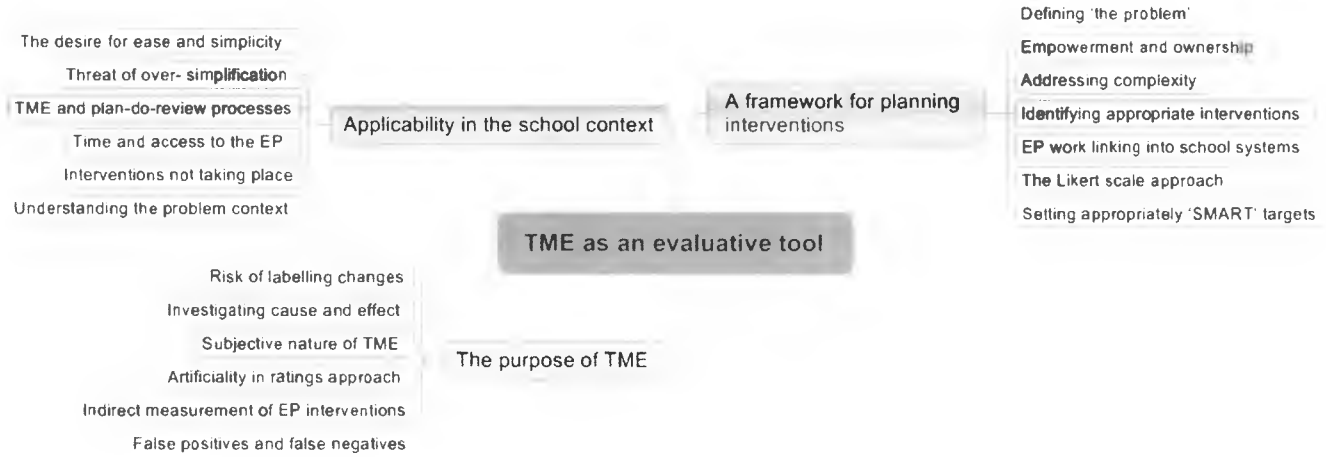
10. **In sum, what do you think the utility of TME is, in light of everything else you've said?**

Appendix 5: Thematic maps

Superordinate theme 1: Professional roles

Superordinate theme 2: Resistance

Superordinate theme 3: TME as an evaluative tool



Appendix 6: Learning checklist

Pupil's Name:	D.O.B:		Year:	
School:	EP:			

Key: **0** = not achieved **1** = achieved with support **2** = achieved independently **3**= secure

	Baseline		Follow-up	
Dates of Assessments:				
Setting (as appropriate):	**Class**	**Ind.**	**Class**	**Ind.**
• Good upright posture				
• Looks up, head held high				
• Sustains good eye contact with adult				
• Sustains good eye contact with peers				
• Confident facial expression				
• Generally open body language				
• Appropriate volume of speaking voice				
• Answers in sentences (not monosyllabic)				
• Appropriately talkative				
• Doesn't try to control session / change agenda				
• Will ask questions				
• Will ask for help/guidance				
• Completes set tasks				
• Perseveres to end of activity / task				
• Listens well to instructions etc				
• Accepts praise				
• Accepts session rules/boundaries				
• Good general attitude				
Total				

Follow up interview:	
• Tell me a bit about how you are finding school at the moment?	
• Is there anything you've noticed about your learning that has changed since we last met?	
• What do you think you are best at in school?	
• Is there anything that you would like to change about your school?	

Appendix 7: Initial consent letter to schools

19th February 2009

Department: **Psychology Service**
Education and Children's Services

Contact Name: Tom Connor
Contact No:
Fax:
Email:
Our Ref:

Dear Headteacher

Re: Trainee Educational Psychologist's Doctoral Thesis

I am a Trainee Educational Psychologist, who will be based within the X Psychology Service until Summer 2010. As part of my training, I will be completing a thesis investigating the impact of Educational Psychology (EP) interventions within the Borough.

During the academic year 2008 – 2009, the X Psychology Service is piloting "Target, Monitoring and Evaluation" (TME) as a means by which to evaluate the impact of interventions in which the EP has involvement. TME is completed in conjunction with teachers and TAs, and involves the setting of specific targets which are subject to review 6 to 8 weeks later. An example form is attached. My thesis will look at the use of TME as a means of investigating outcomes, compared with data obtained using more "traditional" assessment materials.

I will be focusing on children where there is either a concern regarding literacy or behaviour difficulties. It is anticipated that the methods for evaluating change will fit naturally with the day-to-day use of TME as part of the range of tools used by EPs when engaging in teacher and parental consultations. Target setting and follow-up of TME will be completed by the relevant EP and stakeholder.

My role, with your agreement, will be to visit your school and complete short assessments of the child relevant to each TME at the initial and follow-up (review) stages. In the case of literacy, this will involve a short literacy based assessment and, in the case of behaviour, this will involve a short classroom observation. Following the conclusion of data-collecting, interviews with school SENCOs will be sought in order to gauge an understanding of how the use of TME has been valued (or not) in-school.

The research will focus only on those children currently attending mainstream primary schools. Specifically, analysis will involve cases involving children within Key Stage 2 (and, therefore, aged between 7 and 11 years). It is anticipated that data collection will begin in February 2009 and last for approximately two terms, finishing in the summer term (June /

July 2009).

My research has been scrutinized and passed by the Ethics Committee at the Institute of Education. At the point of analysis the data gathered will be anonymous in relation to both the children and the school. Data will be confidential.

I will very much appreciate your support with my project, and am happy to discuss this, or TME, with you in further detail.

Yours sincerely,

Tom Connor
Trainee Educational Psychologist

Appendix 8: YARC outcomes

Baseline and outcome TME measures were compared with outcomes derived from the York Assessment of Reading Comprehension ("YARC"). The results for each YARC component scale are shown in **Table 4.7** below, including baseline and follow-up score, and the total number of cases included in each analysis.

Table 4.7 YARC Outcome Scores

YARC Scale	No. of cases	Mean age equivalent in months	
		Baseline	Follow-up
Reading Accuracy	13	80.7	84.2
Reading Comprehension	13	89.4	92.8
Letter/ Sound Knowledge	14	76.9	78.7
Early Word Reading	14	79.4	80.8
Sound Isolation	14	76.2	79.4
Sound Deletion	14	73.9	78.9

a) Reading Accuracy

Table 4.7 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the reading accuracy condition. A paired samples t-test revealed that these differences were significant ($t(13) = 3.638$ $p < 0.003$).

The baseline and follow-up reading accuracy scores were compared in order to derive a reading accuracy progress score for each case. Further analysis revealed there was no significant correlation between this YARC reading accuracy progress measure and the TME outcome measure (i.e. the mean difference between the follow-up and baseline measures).

b) Reading Comprehension

Table 4.7 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the reading accuracy condition. A paired samples t-test revealed that these differences were not significant ($t(13) = 1.923$ $p > 0.05$).

A YARC reading comprehension progress score was calculated. There was no significant correlation between this and the TME outcome measure.

c) Letter Sound Knowledge

Table 4.7 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the letter sound knowledge condition. A paired samples t-test revealed that these differences were significant ($t(14) = 4.080$ $p < 0.001$).

A YARC letter sound knowledge progress score was calculated. There was no significant correlation between this and the TME outcome measure.

d) Early Word Reading

Table 4.7 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the early word reading condition. A paired samples t-test revealed that these differences were significant ($t(14) = 4.616$ $p < 0.001$).

A YARC early word reading progress score was calculated. There was no significant correlation between this and the TME outcome measure.

e) Sound Isolation

Table 4.7 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the sound isolation condition. A paired samples t-test revealed that these differences were significant ($t(14) = 4.205$ $p < 0.001$).

A YARC sound isolation progress score was calculated. Further analysis revealed there was a significant correlation between this progress measure and the TME outcome measure (i.e. the mean difference between the follow-up and baseline measures), $r(12) = .619$, $p < 0.05$. Therefore, higher Literacy-based TME outcomes were correlated with better performance on the sound isolation task at follow-up.

f) Sound Deletion

Table 4.7 indicates that the mean *follow-up* outcome was higher than the *baseline* rating in the sound deletion condition. A paired samples t-test revealed that these differences were significant ($t(14) = 7.033$ $p < 0.001$).

A YARC sound deletion progress score was calculated. There was no significant correlation between this and the TME outcome measure.

Appendix 9: Superordinate themes by EP and SENCO comments

Superordinate theme 1: Professional roles

Category	Sub-themes	Interviews
The role of the EP	Referral criteria	EP 1,2,7,6,5,8,9 SENCO 1,8,4
	Establishing shared outcomes	EP 1,3,6,5,9 SENCO 3,4,6,8
	Perceptions of the EP role	EP 1,2,3,4,5,6,8 SENCO 2,4,7
	Contrasting approaches to EP case work	EP 2,4,5,7,8,10 SENCO 2,4
	Evidencing EP practice and positive outcomes	EP 1,3,4,5,6,8,10 SENCO 1,4,6
The presence of other factors	Identifying the 'change' variable	EP 2,5,6,9,10 SENCO 2,3,5,6
	Individual professional contributions	EP 3,4,6,7,8,10 SENCO 2,5
	Subjective criteria for identifying progress	EP 2,4,6,7 SENCO 1,4
	Evaluating at EP service level	EP 1,3,4,6,8,9,10 SENCO 1,3,7
	TME and consultation	EP 1,2,4,6,8,9 SENCO 3,4,8
Involving parents and children	Involving parents in evaluative processes	EP 3,4,5,6,7 SENCO 1,5,7
	Managing parental expectation	EP 1,5,6,9 SENCO 4,5,6,7
	Desirability for sharing information	EP 1,4,6,8,10 SENCO 2,3,4
	Involving children in evaluative processes	EP 2,3,4,5,6,9 SENCO 2,3,5,7
	Challenges in involving children	EP 2,4,8,9,10 SENCO 1,2,4,7
	The absent voice of the child	EP 2,6,7,9,10 SENCO 3,4

Superordinate Theme 2 - Resistance

Category	Sub-themes	Interviews
Perceived value of TME	Positive perceptions of TME	EP 1,2,3,4,5,6,9,10 SENCO 1,3,4,7
	Negative perceptions of TME	EP 2,4,7,8 SENCO 1,2,8
	Resistance to change	EP 1,4,7,8,9,10 SENCO 1,4,6,7
	Shared values for evaluation	EP 1,2,5,6,8,10 SENCO 3,5,6
	Goodness of TME fit with effective practice	EP 1,3,6,7,8,9,10 SENCO 5,8
	Change and uncertainty on working practice	EP 2,4,5,7,8 SENCO 2,3,7
Training and confidence	Low confidence	EP 3,4,5,7 SENCO 4,6
	Value of training	EP 1,2,4,5,6,9 SENCO 3,4
	Interactions between EP / SENCO	EP 1,4,5,6,7,9,10 SENCO 1,2,3,6
	Establishing a plan-do-review framework	EP 3,4,5,6,9 SENCO 3,6
	School-based adaptations of TME	EP 2,4,9 SENCO 2,3,4
Establishing an evaluative tool	Suitability of evaluative processes	EP 1,2,4,6 SENCO 2,3,4
	Over-reliance on one evaluation method	EP 1,4,5,6,7,10 SENCO 2,4,5,7
	Target setting approaches	EP 1,5,6,9,10 SENCO 1,2,5,7,8
	Short and longer term change	EP 1,2,4,5,6,7,9 SENCO 3,5,4

Superordinate Theme 3: The practicalities of TME

Category	Sub-themes	Interviews
Applicability in the school context	The desire for ease and simplicity	EP 2,6,8,9,10 SENCO 1,2,4,6,7
	Threat of over- simplification	EP 1,3,6,7,9 SENCO 3,5,8
	TME and plan-do-review processes	EP 3,6,7,9,10 SENCO 2,5,7
	Time and access to the EP	EP 2,4,5,6,7,9 SENCO 1,2,3,4,5,6,8
	Interventions not taking place	EP 4,5,7,8,10 SENCO 1,2,4,6,7
	Understanding the problem context	EP 1,4,6,7,10 SENCO 1,3,5
A framework for planning interventions	Defining 'the problem'	EP 1,2,4,6,9,10 SENCO 2,3,6,7
	Empowerment and ownership	EP 1,6,5,7,10 SENCO 3,5,8
	Addressing complexity	EP 1,3,7,9,10 SENCO 1,5,7,8
	Identifying appropriate interventions	EP 2,4,6,9 SENCO 1,2,4,6
	EP work linking into school systems	EP 2,3,6,10 SENCO 2,3,4,6,7
	The Likert scale approach	EP 1,2,4,5,6,10 SENCO 1,2,4,5,7,8
	Setting appropriately 'SMART' targets	EP 1,2,5,6,7,8,9,10 SENCO 3,5,6,7
The purpose of TME	Risk of labelling changes	EP 2,4,6,7,9,10 SENCO 1,3,4,6,7,8
	Investigating cause and effect	EP 1,2,5,9,10 SENCO 2,3,5
	Subjective nature of TME	EP 1,2,4,6,7,8,9 SENCO 1,2,4,7
	Artificiality in ratings approach	EP 1,4,6,7 SENCO 2,5,6
	Indirect measurement of EP interventions	EP 1,2,4,5,6,7,8,9 SENCO 1,3,4,6,7
	False positives and false negatives	EP 2,4,5,6,7 SENCO 2,3,5