

**OCCASIONAL PAPER No. 1
1996**

Quality Evaluation Of Early Childhood Education Programmes

Pamela Cubey and Carmen Dalli

Series Editor: Carmen Dalli

Institute for Early Childhood Studies
Victoria University of Wellington
PO Box 600, Wellington, New Zealand

©1996 Copyright is held by individual writers over their own work
Publishing rights are held by the Institute for Early Childhood Studies

ISSN: 1174-1546
Occasional Paper Series No. 1
[Series Editor: Carmen Dalli]

ISBN: 0-475-20050-0
Quality Evaluation of Early Childhood Education Programmes

CONTENTS

Introduction	3
The terms "evaluation" and "assessment"	4
Why assess? Why evaluate?	6
Evaluation and the early childhood curriculum	9
The evaluation process, reflective practice and quality early childhood education	12
Assessment and evaluation practices	14
Testing and conventional approaches	14
Alternative approaches	17
Methods and instruments of evaluation	18
Constraints on evaluation of early childhood programmes	20
Concluding summary	21
References	23

Introduction

In 1989 the New Zealand Council for Educational Research organised a seminar on Assessment in Early Childhood Education. This was one of the first times that the topic of assessment and evaluation received formal airing by the early childhood community. Until then, the generally accepted wisdom had been that assessment in early childhood education was inappropriate because of the very young age of the children who use this service and the great variability in developmental progress at this age.

The 1990s saw a decided increase in interest in assessment and evaluation. Curricular change, and with it, assessment and evaluation, became part of the national policy agenda for all parts of the educational sector.

In early childhood education, the development of *Te Whaariki*¹ (Ministry of Education, 1996), the early childhood curriculum, has been accompanied by an increasing awareness that just as the early childhood curriculum is quite distinct from the more formal and subject-oriented one characteristic of the schooling system, so too there are methods of assessment and evaluation in early childhood that are more appropriate than others (Carr, 1989; Smith 1989).

It has also become more accepted that assessment and evaluation are a necessary part of an educational service that aspires to be a high quality provision. New Zealand early childhood centres are audited by the Education Review Office for compliance with regulations and this is seen as one evaluative measure that is related to the provision of a quality service; discussions, however, are ongoing about other possible ways in which quality of provision might be assessed, evaluated and /or assured.

This document is intended to contribute to the current discussion by

- clarifying the terms "evaluation" and "assessment"
- focussing on evaluation of early childhood programmes and reviewing recent New Zealand and other literature on this topic
- overviewing the rationale behind evaluation and assessment

¹ *Te Whaariki* is the New Zealand early childhood curriculum developed after extensive consultation with the early childhood community in the early part of the 1990s and officially launched at the end of June 1996.

- considering the relationship between the curriculum and evaluation, and the relationship between evaluation, reflective practice and quality early childhood education
- presenting some approaches to assessment and evaluation
- discussing some possible constraints on evaluation practices.

The terms "evaluation" and "assessment"

The terms "evaluation" and "assessment" often appear side by side. Indeed, some writers, especially American ones, use the terms interchangeably so that it is not always easy to see the boundaries between the two. For example, in a paper on assessment in early childhood programs, Leavitt and Eheart (1991) talk about an assessment process which consists of four closely related and ongoing processes: acquiring information from parents; recording observations of children at play and in daily routines and interactions; developing a comprehensive assessment of each child; and applying ongoing observations and assessments in curriculum planning, implementation and evaluation. There is no clear distinction made in this work between assessment and evaluation.

In other writings, evaluation is seen as an aspect of the broader concept of assessment, while in yet other work, it is assessment that is seen as a subset of a broader endeavour called "evaluation". An example of the former type of writing is the National Association for the Education of Young Children (NAEYC, 1991) document: *Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 Through 8*, which states that "assessment encompasses the many forms of evaluation available to educational decision makers" (p21).

The work of Hurst and Lally (1992), on the other hand, is an example of how the term "assessment" is used to refer to an educational endeavour that is narrower in scope than the more encompassing notion of "evaluation"; while they do not make explicit the distinction between assessment and evaluation, they use the term "assessment" to refer to a specific procedure within a broader process of evaluation. They state that "assessment of children at any stage gives teachers information about the effectiveness of educational processes and procedures. This knowledge can then be applied to the working out of suitable provision for the future (the curriculum) and to the evaluation of educational ideas through critical analysis of their effects on learners" (p46).

In this view, then, the assessment (in this case of children) procedure yields information on outcomes. This information can then feed into the wider process of evaluation. An additional implication of this statement is a view of evaluation as a process that judges worth or value and that leads to action to improve provision.

In a discussion on ways of assessing children and the curriculum, Genishi (1992) uses the following quotation from a conversation at a conference to clarify the confusion that arises from the interchangeable use of the terms assessment and evaluation: "assessment strategies tell you 'what is'. Documentation provides proof of 'what is' and evaluation tells you 'what is in light of what should be...'. Too often we use the words interchangeably and then we get into trouble "(p186).

Other early childhood writers share Genishi's (1992) understanding of evaluation. Faragher and McNaughton (1990) and Gullo (1994) see evaluation as a process that judges or assesses the worth of educational programmes, projects, materials or techniques and involves deciding what is worthwhile or valuable. Kamii (1985) adds the notion that evaluation refers to finding out about the relative effectiveness of alternative approaches to early childhood education.

In a somewhat more political vein, Van der Eyken (1992) sees evaluation as a process which helps practitioners say more clearly what it is they are doing and the nature of the issues being confronted. Bartholemew and Bruce (1993) simply describe evaluation as "what we have offered, are offering and will offer the child next"(p1). More expansively, Hurst (1991) makes the same point when defining evaluation as consisting of:

collecting evidence, reflecting on it, noting particular patterns as they emerge and trying to see what principles of good practice are involved. From this sequence there will emerge certain issues that are seen to be of importance. These issues provide the criteria for decisions about what is the best way to do things (p105).

This last quotation also links evaluation with a principle which is much upheld in professional development programmes: good quality educational practice is reflective practice based on evaluation. Reflective practice is seen to lead to improvement in the quality of the educational provision.

In the New Zealand context, Carr & May (1993b), two of the writers of the early childhood curriculum guidelines, the draft version of *Te Whaariki* (Ministry of Education, 1996), have been strong advocates of this view of evaluation. They argue that the fourth guiding principle articulated in *Te Whaariki*, ie 'responsive and reciprocal' relationships, means that for a curriculum to be responsive, assessment and evaluation must act in a reciprocal way with the aims and goals of the curriculum. Consequently, the curriculum must be evaluated in terms of the aims and goals that make it up. They also note that "assessment and evaluation must start from the aims and goals that make up the curriculum. Otherwise, the assessment procedures will determine the curriculum" (p141).

This view is upheld in the final version of *Te Whaariki* (Ministry of Education, 1996). The document also provides very clear statements on the purposes of evaluation and assessment:

The purpose of evaluation is to make informed judgments about the quality and effectiveness of the programme.....Evaluative procedures emphasise the quality of provision and make use of all the forms of assessment that can be carried out by both adults and children. Assessment of children's learning and development will be part of the information needed to evaluate the programme.....

The purpose of assessment is to give useful information about children's learning and development to the adults providing the programme and to children and their families (p29).

This short discussion of how the terms assessment and evaluation are used in the literature has highlighted two major themes. The first theme is the distinction that exists between evaluation and assessment: evaluation is strongly linked to gauging effectiveness, worth or value of a process or programme; assessment, on the other hand, is associated with measuring outcomes.

The second theme is the connection that exists between evaluation and the provision of a high quality educational service. Both themes are threaded through the discussion in the following sections of this paper.

Why assess? Why evaluate?

An answer to these two questions has already emerged from the preceding section: assessment is carried out for information on the outcomes of specific actions or

programmes and evaluation provides information about their value. It is useful to explore some other reasons behind these processes.

In a book on the design and evaluation of school programs, Eisner (1985a) talks about the many functions of evaluation and singles out five as "especially important" (p192). Eisner's five functions are:

(1) to diagnose difficulties with the curriculum, in the teaching that is occurring or in the student's learning so that an educational "treatment" might be found;

(2) to provide a feedback mechanism for curriculum revision and thus educational improvement. This function of evaluation is seen to "diminish(es) the tendency to use evaluation practices simply as a means of classifying students, rewarding them, or selecting the able from the rest" (p195); it also shifts the full responsibility for performance away from the student and closer to the school or educational setting. Eisner adds that "when student performance is viewed as an index of program effectiveness, the likelihood of curriculum improvement increases and a major contribution is made toward improving the quality of education"(p 196).

(3) to compare programs, teaching and other aspects of the educational process to make available evidence that may be used in decision-making on educational policies;

(4) to anticipate or identify educational needs; this is a form of needs assessment or needs analysis so that the existence of a need may be established and a proposal advanced about how the need may be met;

and

(5) to determine if objectives have been achieved; objectives are defined by Eisner as "the criteria for determining whether the program has been effective" and evaluation is "the means of collecting the data and analysing them"(p198).

Another important aspect of Eisner's (1985a) work is his distinction between the "three important subject matters" for evaluation: "the curriculum itself, the teaching that is provided and the outcomes that are realised" (p 200). He further divides outcomes into three kinds: subject-specific outcomes, student-specific outcomes and teaching-specific ones (pp205-208). While Eisner's work relates to evaluation in a school context, his theoretical discussions on evaluation also have relevance for the early childhood setting. Various other writers deal only with one or other of these three subject matters but Eisner provides a rationale for all three.

Speaking specifically about evaluation in early childhood contexts, Hurst (1991) suggests that the need for the information provided by assessment and evaluation arises

out of an educational context in which accountability has become an integral part of educators' terms of reference. She writes that "today, with increasing pressure for accountability, all who are concerned with the care and education of young children are aware that some form of monitoring is called for ...and the actual practitioners need to play a central role" (p101).

Later Hurst adds that "what unites the three processes of curriculum development, assessment and evaluation is the concern for the progress and development of the individual child " (p105).

These statements clearly link assessment and evaluation with the need to integrate evaluation processes into the teacher's normal practice; evaluation is also linked with the enhancement of children's educational experience.

Van der Eyken (1992), focussing on the evaluation of educational projects, makes a similar point about the aim of evaluation. Evaluation should be to "enhance, not denigrate" and should always be undertaken in a spirit of "valuing a project" (p16).

It is important to note, however, that these functions of evaluation cannot simply be assumed to occur for each evaluation event. As Eisner (1985a) has argued, the way that evaluation is carried out, the conception of education that underlies the focus of evaluation, and the methods used to conduct evaluation, all influence the consequences of evaluation. This is clear in Hoot, Bartkowiak, Goupil, Mercado and Panepento (1993) who note the inadequacy of teacher evaluation instruments used by administrators to assess effective teaching of early childhood public school teachers in seventeen school districts in Western New York. Hoot et al note that "teachers often view evaluation by administrators as a rather superficial experience and of little use to them 'as practitioners interested in their professional development' " (p 5). They argue that the major purpose of evaluation should be the "improvement of the professional activities of those in the field" (p 12) and suggest that this is more possible when the evaluation instruments used are based on guidelines about which the profession is in agreement.

The expectation that educational practice might be improved by evaluation and assessment is a favourite theme of Hurst and Lally who together comment (Hurst and Lally, 1992) that assessment is "closely related to the evolution of ideas about educational quality as well as the development of appropriate curricula" (p 61). In addition, they state that assessment is "important because it offers ways in which to construct an educational critique of proposals and procedures" (p 46)

Hurst's (1991) work on planning for early learning makes a further interesting point on the aims of evaluation. She suggests that when staff evaluations are being carried out, the aim should be broader than just to find out how well the staff are doing what they set out to do. She argues that evaluations will enable staff to ask questions about what it is they are setting out to do and thus to provide a critique of practice, because: "unless evaluation fuels and guides this kind of critique it is only capable of endless rubber stamping the initial plans of the evaluator "(p 118).

Hurst's argument here is clearly similar to Hoot et al (1993); in addition, she argues that evaluations should also help staff to notice what it is they are doing which they did not set out to do. In this way it may be possible for the hidden curriculum to become more easily identified.

It would appear therefore that the purposes of assessment and evaluation could be summarised as follows:

- to diagnose
- to revise curricula
- to compare
- to anticipate educational needs
- to determine if objectives have been achieved
- to satisfy demands for accountability
- to improve educational outcomes for children
- to critique professional practice
- to enhance professional practice
- to value a project or a programme
- to uncover the hidden curriculum.

Evaluation and the Early Childhood Curriculum

In Aotearoa New Zealand the early childhood curriculum has been defined as "the sum total of the experiences, activities, and events, whether direct or indirect, which occur within an environment designed to foster children's learning and development" (Ministry of Education, 1996, p10).

Anne Stonehouse's (1991) definition is equally inclusive; she sees the curriculum happening "all the time that at least one child is present"(p 101).

The inclusiveness of these statements should not be interpreted as a *carte blanche* for any activity or experience or for a *laissez faire* attitude to the way the curriculum develops. Carr & May (1993a) are quite clear that there is an "appropriate" curriculum for early childhood as well as an "appropriate" way of going about planning it. They state:

the right way to go about curriculum planning at any level is to start with the appropriate curriculum and then to look at how to assess or evaluate it. And then to keep a lookout to ensure that the assessment does not skew the curriculum offered (p10).

As has already been noted, Carr and May (1993b) are firmly of the view that assessment and evaluation should start from the aims and goals that make up the curriculum ; the reverse relationship would sabotage the curriculum. The final version of *Te Whaariki* (Ministry of Education, 1996) has remained true to this statement. It states that "it is essential that assessment and evaluation are based on the goals of each strand of the curriculum and that the principles of the curriculum are always applied" (p29).

This view is strongly advocated by many writers (eg. Gura, 1992; Smith, 1989; Eisner, 1985a; Kamii, 1985; Honig, 1994 ; N.A.E.Y.C., 1991; Hurst, 1991; Hurst & Lally, 1992; Gullo, 1994). Alice Honig (1994) makes the additional point that : "when evaluative components are carefully built in with the philosophy and its implementation they can often serve as a powerful adjunct to enhance the quality of the service provision" (p 1).

Sitting comfortably with the above are the assertions of Blenkin & Kelly (1987) that "educational planning must begin from a consideration of the development of the child who is the subject of that planning" (p 2) and of the N.A.E.Y.C. (1991) that curriculum and assessment provisions must be based on knowledge of what is age and individually appropriate (p 28). *Te Whaariki* (Ministry of Education, 1996) adds the cautionary note that the "needs of the children, not assessment procedures, should determine the curriculum"(p 29).

It is clear from the above that evaluation and the curriculum are seen to be causally linked and that there is a definite view on the way the direction of causality should flow,

namely downwards from the goals of the curriculum to its implementation and onto the evaluation which itself should start from the original goals.

Picking up on a different point, Van der Eyken (1992) notes that "projects that have vague and unclear goals tend to produce diffuse and unfocussed evaluations. Good evaluation depends on clarity. This is not the same as simplicity" (p 3).

This statement highlights the issue of what constitutes "good" or "valid" evaluation. Gullo's (1994) answer is that if evaluation is to be valid there must be a close relationship between it and the curriculum, with the particular evaluation or assessment content and procedure reflecting the curriculum content and the instructional strategies that are utilised. He further observes that just as curriculum experiences are beneficial to different children in different ways, there may be some curriculum experiences that are effective for no children. Here, evaluation can be a useful instrument in making curriculum adjustments.

Like Carr and May (1993b) and Meade (1993), Gullo (1994) also maintains that evaluation should not determine the type of curriculum content or strategies.

The concern that assessment and evaluation might become the proverbial "tail that wags the dog" is not the only danger to beware of. Lally (1991) voices a concern of a different nature:

this (evaluation) process of finding out, using information to make plans, monitoring plans in action and sharing perceptions lies at the heart of the responsive nursery curriculum...unless they (the teachers) are in control of the process of monitoring and evaluation, they run the risk of having their work judged by other people's standards rather than by their own (p 119).

This statement of Lally's is a timely reminder of the "political function" of evaluation which Blenkin & Kelly (1987) predict in their much-referred-to publication *Early Childhood Education: A Developmental Curriculum*. They state:

in the future, evaluations will be required to fulfil two purposes, one a political and the other a professional function. The professional purpose is placed at risk if the external accountability is strengthened, especially if the latter runs counter to the former (p 29).

Blenkin and Kelly's argument highlights the need for evaluations to be determined by the teachers themselves (see also Hurst, 1991) and that they develop competence in this field. This competence has a political function because it is likely, for example, to enable teachers to be more articulate about their goals and about what they perceive their role to be as the providers of educational services. The professional function of evaluation, that is the potential for self-development that comes from an evaluation that leads one to reflect on one's practice, would be jeopardised if teachers were not allowed to control the process of evaluation themselves.

The evaluation process , reflective practice and quality early childhood education

The preceding discussion on the relationship between evaluation and the curriculum noted that there is agreement in the literature that the two are inextricably connected; evaluation has to be based on the aims and goals of the curriculum. Faragher & MacNaughton (1990) take this point a step further; they state that evaluation "is a critical part of a quality programme for young children. It is a vital and ongoing part of the programme planning process. It is both the final step in planning and the beginning of the planning process" (p 143).

This statement suggests that once evaluation has occurred, the results of the evaluation should re-start the curriculum planning process thus creating a cyclical pattern of planning, acting, evaluating and yet more planning, acting and evaluating. The writers are clear that one cannot run a quality early childhood programme without this continuous process.

Other writers (eg Blenkin & Kelly, 1992; McCrea, 1989 ; Duckworth, 1991; Killion & Todnem,1991; Haggerty & Hubbard, 1993) also refer to the valuable place of evaluation in the cyclical form of programme planning while Gullo (1994) emphasises the need for assessment and evaluation procedures to "describe the progress of children over time....a description of where children are at at any given moment within some learning sequence continuum" (p 22).

Hurst (1991) also views evaluation as a cyclical pattern of processes, with the centre staff being able to reflect, at each part of the cycle, on the evaluation criteria they have established and then being able to take appropriate action .

Several other writers refer to the importance of reflective evaluation and practice (eg Gura, 1992 and 1993; Blenkin & Kelly, 1987 and 1992; Duff, Brown & Van Scoy, 1995; Haggerty & Hubbard, 1994; Meade, 1985; Eisner, 1985a; Zimiles, 1993; Duckworth, 1991; Wellington, 1991; Killion & Todnem, 1991; Hoot et al, 1993; Drummond, Rouse & Pugh, 1992; Hurst, 1991; Genishi, 1992; and Carr & May, 1993b). Many writers link reflective practice and evaluation to true professionalism and to teachers' professional development; Blenkin and Kelly (1987) state that "this two-way process of translating ideas into practice and reflecting upon the results is at the heart of genuine professional work" (p 25).

Eisner (1985b) puts his hope for the improvement of education in the ability of teachers to become reflective practitioners:

I start with the assumption that the improvement of education will result not so much from attempting to find scientific methods that can be applied universally...but rather from enabling teachers and others engaged in education to improve their ability to see and think about what they do (p 104).

Similarly, Genishi (1992) writes of the concept of "teacher-as-learner and teacher-as-researcher, asking genuine questions...reflecting on the daily events and dramas...and raising questions about their policies and practices that affect children and teachers" (p 195).

A further point that is made is that while reflection can be done alone, in early childhood settings, where staff work as a team, reflective evaluation can benefit all staff when they are all involved in discussion about what they have observed and documented (eg Bruce, 1987; Gura, 1992 & 1993; Meade, 1985; Blenkin & Kelly, 1992; Honig, 1994; Van der Eyken, 1992; Lally, 1991; Hurst, 1991; and Foote, Stafford & Cuffaro; 1992). Van der Eyken (1992) says that "evaluation is a collaborative exercise. It belongs to everyone concerned and only if people are involved will its outcomes be accepted and valued and will it cease to be a marginal activity" (p3). Later he adds that:

A project should retain control of and involvement in its own evaluation. At the end the staff should be able to say, 'we don't necessarily agree with all the report but we think it is fair and everyone has had a chance to comment, and has been listened to' (p 28).

Writing within the context of developing the New Zealand early childhood curriculum guidelines, Carr and May (1993b) voice similar hopes about the evaluation that will accompany the implementation of *Te Whaariki*:

It may be that one of the greatest contributions of a national curriculum to improving the quality of early childhood programmes is the discussion and reflection that accompanies its development...Another measure of its value will be whether discussion and reflection continue to contribute to the national curriculum to create changing patterns for individual whaariki and to suggest reviews of the guidelines (p 152).

Clearly, collaborative evaluation that feeds back into the planning of the curriculum is seen as an instrument for enhancing the quality of early childhood programmes. This view is consistent with Torrance's (1991) who says that when assessment is used more to diagnose, guide and support and less to specify, measure and certify, it is related to improving quality. Reflective practice of this kind is also considered to be at the heart of professional practice.

Assessment and evaluation practices

The literature reviewed above presents a strong argument for a close match between the curriculum and evaluation. Since assessment is such an integral part of the evaluation process, it is crucial to be aware of both the advantages and the potential dangers associated with the various assessment approaches. Focussing on assessment instruments and strategies, Gullo (1994) notes how important it is that the curriculum should determine which instruments are used, while Blenkin and Kelly (1987) advocate procedures that will promote rather than restrain staff's work as educators.

Testing and conventional approaches

It has been noted already that the New Zealand early childhood community has traditionally been wary of the conventional testing approach to assessment, and early childhood scholars have long argued that any assessment in early childhood should be "appropriate" (eg Carr, 1989; Smith 1989; Farquhar, 1989; and Meade, 1990). A major theme they emphasise is that "appropriate assessment" is assessment that is linked to action: appropriate assessment is assessment that not only diagnoses difficulties but works to improve educational outcomes for children through, for example, revising the curriculum.

Smith (1989) identifies three possible purposes for assessment during the early childhood years: assessment for accountability, assessment for categorisation and assessment for planning and monitoring. She points out that both assessment for accountability and assessment solely for categorisation can be very dangerous.

Smith does not reject assessment for accountability but points out that the pressures of external measures could well result in the adoption of "short term goals in achievement ... at the expense of real understanding and ongoing motivation" (p39).

Smith calls for "assessment procedures that are based on sound educational principles" and "communicated to parents and the wider community" (p40).

Smith's objection to assessment for categorisation is that it is effectively a labelling process that can easily result in self-fulfilling prophecies. Research on the self-fulfilling prophecy shows that, generally, this process is allowed to take its course despite the fact that it is now well-accepted that development is a discontinuous process and that life experiences can greatly modify developmental outcomes.

Other scholars writing from within early childhood traditions that are similar though different to the New Zealand one, also express concern about the dangers of using standardised tests to assess young children and about their lack of reliability and validity (eg Bartholemew & Bruce, 1993; Blenkin & Kelly, 1992; Genishi, 1992; Gullo, 1994; Hurst, 1991; Kamii, 1985; NAEYC, 1991; and Schweinhart, 1993).

Kamii (1985) feels that conventional assessment methods overlook the necessity for evaluating the goals and objectives of a curriculum and may also be used to define the educational objectives.

Farquhar (1989), like Smith (1989) and Kamii (1985) links "appropriate assessment" to educational goals and philosophy; she also links it to action: what is needed is assessment "that is formative" in nature, tied in with society's goals for early childhood education and care and ... relevant to the early childhood field" (p 59). There is obvious agreement here with Carr's (1989) view that there is "no sense in finding out that 50 percent of the children can't label 10 colours if we are not planning to do something about it" (p 18).

Clearly then, appropriate assessment in early childhood is seen by New Zealand writers to be tied to planning and monitoring and to "be driven by an intention to change a

child's behaviour not to classify or compare with a normative model" (Smith, 1989 p 44). In analysis, this view is consistent with the shift away from mathematical and statistical models of assessment and towards educational and psychological ones which has taken place in the assessment for quality debate in the United States (Torrance, 1991).

As Eisner (1985a; 1985b) has pointed out, the methodology of educational evaluation has been dominated by scientific assumptions and one of the negative consequences of this is that the term *evaluation* has been reduced to *measurement* (1985a p14). He argues that

although each term is entirely independent - that is, one can evaluate without measuring and one can measure without evaluating (one can, of course, do both) - the belief that one must measure in order to evaluate is widespread. When this occurs, the fields that are most amenable to measurement are measured and those that are difficult to measure are neglected. What is measured then is emphasised in school programs because measurement becomes the procedure through which educational quality is determined (pp 14-15).

The danger that evaluation will lead the curriculum and thus become "the tail that wags the dog" is also highlighted by Eisner when he warns:

educational practices based on a scientific model too often become not a tool for improving the quality of teaching and learning but rather an impediment to such ends. Students and teachers alike gear up to take tests, even though none of them believes those tests to be intrinsically important or that the tests really assess much of what has been learned and taught in schools (p15).

In another publication (Eisner, 1985b), he adds

It is not the goals that drive the system as much as it is the way in which the system is to be evaluated: the form of the evaluation and the content to which it attends become the operational goals of the system (pp 224).

Another aspect of Eisner's argument is that scientific assumptions and the methods they yield do not exhaust the ways in which we come to know. Indeed, "their exclusive use has led to a limited and parochial conception of how educational evaluation can proceed" (p117). Eisner argues that the reason the testing approach to evaluation has become so influential in shaping practice is "because evaluation results have a public

status that neither curriculum nor teaching have. They can be inspected and appraised by the public." (p223).

As an alternative to the scientific approach to evaluation, Eisner describes two concepts: educational connoisseurship and educational criticism, which are embedded in an artistic rather than in a scientific tradition. He argues that these concepts could "provide a needed complement to the scientific approaches used today" and gives suggestions as to how they may be applied as well as identifies some "unanswered questions and potential problems" (pp103-117). His argument is not for a replacement of existing methods but for the need for an expansion of methods.

Alternative Approaches

In Genishi's (1992) *Ways of Assessing Children and Curriculum*, teachers and researchers in six different settings provide accounts of successful alternative approaches to assessment, all in classrooms regarded as developmentally appropriate, where "teacher- and child-based assessment go hand-in-glove with teacher-and child-based curriculum"(p 5). These teachers mostly used informal ways of observing and documenting development and learning. A number of writers (eg Bartholemew & Bruce,1993; Blenkin & Kelly, 1992; Hurst & Lally,1992 and Kamii, 1985) are in favour of using alternative assessment in lieu of standardised testing for young children. Gullo (1994) opines that alternative assessment procedures have some distinct advantages over conventional ones and provides a useful summary of the differences between alternative assessment and conventional assessment procedures. Advantages of alternative assessment procedures include the usefulness of the information gained, the focus on children's individual development and the link that is maintained between the curriculum, the learner and evaluation. Gullo states that in alternative assessment procedures:

There is a close match between curricular goals and assessment outcomes. Thus, the resulting information is relevant for further curriculum development and modification..... by using actual curricular activities as the evaluation means, a broader measure of curricular effectiveness can be assessed (p 86).

Gullo also adds that:"one of the advantages of using an alternative assessment approach is that it does not disrupt the process of curriculum implementation" (p 90).

In a study of evaluation practices of model early childhood programmes, Slavenas (1993) found that open-ended, process-oriented methods, such as observations and interviews, were used more often than predetermined methods such as questionnaires and check sheets. This is in agreement with the view advocated by numerous writers (eg Blenkin & Kelly, 1987; Smith, 1989; Carr & May, 1993b; Haggerty & Hubbard, 1994; Hurst, 1992; Hurst & Lally, 1992; Koralek, 1993; NAEYC, 1991; DeStefano, 1991; Farquhar, 1993; Harms & Clifford, 1993; Honig, 1994; and Leavitt & Eheart, 1991) that alternative assessment fits comfortably with the developmental approach to curriculum.

When the purpose of evaluation is to determine the effectiveness of an ongoing programme so that it may be modified and enhanced, it is often referred to as *formative* evaluation.

Gullo (1994) refers to specific instruments designed by Derman-Sparks (1989), Schweinhart (1993) and Harms and Clifford (1993) to evaluate different programme components in a formative evaluation. He believes that: "formative evaluation should take place on a continuing basis during the implementation life of the program rather than at the end...when modifications would not be beneficial to those completing the program" (p 19).

In Gullo's view, at the end of the educational experience, a *summative* evaluation should occur looking at the overall effectiveness of the experience. He also believes that formative and summative evaluation should go hand in hand.

Bruce (1987) sees the need for both formative and summative record keeping, with the latter being dependent on the former; she believes this "supports the interactionist approach to early childhood education" (p 167).

DeStefano, Maude, Crews and Mabry, 1991; Farquhar, 1993; Harms & Clifford, 1993; Honig, 1994; Hurst, 1991; Leavitt and Eheart, 1991 and Phillips and Howes, 1987 advocate a flexible approach to evaluation, which means, at times, using several different assessment methods simultaneously to better meet the needs of the wide variety of educational settings. This is in line with Meade's (1991) call for different cultural blueprints and with the approach taken in *Te Whaariki* (Ministry of Education, 1996).

Haggerty (1996) refers to Dadds' (1991 cited in Dadds, 1993) "lack of closure" model in which practice is looked at in the light of theory and theory looked at in the light of practice. Haggerty believes that Dadds' is a model that is implicit in the *Te Whaariki*

curriculum development project: in the same way that *Te Whaariki* is a reciprocal, interactive, empowering, holistic curriculum so it is stated (Carr, 1992c; Carr & May, 1993b) that its evaluation should be reciprocal, interactive, empowering and holistic as well as being ecological to reflect the wider community .

Smith (1989) also advocates an holistic approach to evaluation, arguing, like many others (eg. Bruce, 1987; Young, n.d.; Katz, 1994; Hurst, 1991; Ministry of Education, 1993), that for evaluation to be holistic, parents as well as the wider community need be involved in the process.

It is clear then that, as both Smith (1989) and Meade (1993) have pointed out, formulating assessment policies to accompany curricula is a complex task and will take time because of the "complexity of thinking needed to be worked through" (Meade, 1993, p56). This comment also applies to the formulation of appropriate evaluation policies which do not disrupt the process of curriculum implementation.

Methods and instruments of evaluation

There is consensus among writers referred to in this review that the primary means for effective evaluation of early childhood programmes are competent observation and record keeping. Various approaches to these methods are possible including the use of narrative, anecdotal records, journals, profiles, portfolios of children's work, photographs and video recordings. Genishi (1992) also makes a case for the value of participant observation saying that: "systematic observation keeps every child in the foreground for some period of time, so that all children become part of remembered scenes" (p 196).

These methods represent the tools of continuous assessment and documentation which make overall evaluation easier. Table 1 below presents the sources of some of the more useful instruments for observation and record-keeping together with other useful information about assessment and evaluation methods.

<i>Type of information on Instruments</i>	<i>Source</i>
<p>General information</p> <ul style="list-style-type: none"> • Advantages and disadvantages of different instruments • Glossary of assessment instruments • Description of different instruments • Choosing evaluative measures 	<p>Harms & Clifford (1993) Gullo (1994) Slavenas (1993, pp38-39) van der Eyken (1992)</p>
<p>Observation and Record-keeping</p> <ul style="list-style-type: none"> • A variety of approaches • Record-keeping and assessment • Checklist for evaluating the environment from the child's viewpoint • Evaluative questions from a child's viewpoint • Questions to evaluate children's development of autonomy in learning and in their relationship with adults and other children • Infant-toddler environment rating scale 	<p>Bartholemew & Bruce (1993) Lally (1991)</p> <p>Harms (1981)</p> <p>Katz (1995)</p> <p>Kamii (1985:6-7) Harms, Cryer & Clifford (1980)</p>
<p>Evaluation of Quality Provision</p> <ul style="list-style-type: none"> • Checklists for evaluation thedevelopmental appropriateness of learning experiences, strategies, organisation and examples of evaluation in practice • Examples if reflective questions illustrating how the aims and goals of <i>Te Whaariki</i> may be illustrated • Evaluation approach in searching for exemplary paractice in early childhood education • Ways in which to view quality • Developing measures to assess quality daycare, including home-based care • Instruments for identifying key components of high quality educational settings taking account of contextual influences on settings • Questions for evaluation a programme's assessment procedures • Instrument used in NAEYC accreditation process 	<p>Faragher & MacNaughton (1990:147-152)</p> <p>Carr & May (1993)</p> <p>De Stefano et al (1991) Phillips & howes (1987, p3)</p> <p>Mooney et al (1994)</p> <p>Harms & Clifford (1993 , Chapter 30) Rossbach, Clifford & Harms (1991)</p> <p>NAEYC (1991, pp34-35)</p> <p>Bredenkamp (1989)</p>

Table 1: Instruments for evaluation

Constraints on evaluation of early childhood programmes

It has already been noted that effective evaluation is a complex process (Smith, 1989; Meade, 1993); in the New Zealand context where a tradition of assessment and evaluation does not exist in the early childhood sector, this may mean that it might be some time before the early childhood community is fully at ease with evaluation practices.

More pragmatic constraints on developing effective evaluation practices also need to be noted. One of these is the difficulty of finding the time and opportunity which are needed for discussion and reflection. Many writers (eg Gura, 1992; Meade, 1985 & 1993; Mooney et al, 1994; Honig, 1994; Lally, 1991; Waters et al, 1992) have noted that early childhood staff often do not have provision for this .

Staff also need to be clear about their intentions and how they will act upon them and this means knowing about child development and the curriculum (Bruce, 1987; Blenkin & Kelly, 1992; Honig, 1994; Smith, 1989) as well as about evaluation . This is likely to require further training and support for early childhood staff (Blenkin & Kelly,1987; Eisner, 1985; Duckworth, 1991; Honig, 1994; Hoot et al, 1993) because, as Smith (1989) notes, "a thoroughly professional attitude is needed which is unlikely to be acquired casually or informally" (p 31).

A further constraint is pointed out by Hurst (1991) who makes the important point that "evaluation can be seen as rather a threatening experience" (p1 18). She suggests that it should therefore "have regard for the sensitivities of those involved and should be undertaken in a professional spirit with the intention of furthering the practice and development of staff" (p1 18).

Despite these constraints, however, both Hurst(1991) and Smith (1989) as well as Meade (1993) indicate that it is politically imperative that teachers do not eschew evaluations. Hurst points out that one of the advantages of evaluation, particularly selfevaluation that is "under the control of staff ...(is that) it enables (them) to be better able to withstand uninformed criticism"(p1 18). Speaking about assessment, Smith warns that if teachers do not pick up the challenge of being in control of assessment themselves "other 'experts' such as psychologists, health professionals and ministry officials will take away their function from them with possible harmful consequences" (p 31). There is agreement among early childhood scholars that this prospect would not be an appealing one.

Concluding summary

This paper has presented a review of recent literature on the evaluation of early childhood programmes. A number of issues around this topic have been highlighted; in summary, they are as follows :

- 1 There has been confusion in the literature on the difference between the terms "evaluation" and "assessment". Evaluation refers to a process that judges the worth or value of an educational experience; assessment measures the outcome/s of that experience.
- 2 Evaluation has many purposes which include the need to diagnose difficulties, to revise curricula, to compare programmes/ teaching/ processes, to anticipate educational needs and the need to determine if objectives have been achieved. Other purposes of evaluation include reasons of accountability, critiquing professional practice, valuing a programme, improving outcomes for children and enhancing the quality of the educational programme.
- 3 Evaluation and the curriculum are inextricably connected: the evaluation of the curriculum must reflect and be consistent with the philosophy and values that underlie the curriculum. It is the goals of the curriculum that should determine assessment and evaluation and not vice versa.
- 4 Evaluation is critical for a quality early childhood programme. Evaluation is seen as an ongoing part of the programme-planning process; without evaluation there can be no guarantee that the programme meets the aims and goals of the curriculum.
- 5 Evaluation is strongly related to reflective practice which is considered to be best quality professional practice.
- 6 In early childhood settings, alternative forms of assessment are preferred over traditional forms, such as standardised tests; alternative forms are seen to be less disruptive of the process of curriculum implementation. Observation and record keeping are two of the preferred alternative methods.
- 7 It is vital that early childhood staff are involved in evaluating their programmes to ensure that inappropriate evaluation methods are not forced upon them from

outside sources. This will require training and support so that competence and confidence is built among teachers.

REFERENCES

- Balageur, I., Mestres, J. & Penn, H. (n.d.) *Quality in services for young children* . Discussion paper of the European Commission Childcare Network.
- Bartholemew, L. & Bruce, T. (1993). *Getting to Know You: A Guide to Record Keeping in Early Childhood Education and Care*. London: Hodder and Stoughton.
- Blenkin, G. & Kelly, A.V. (Eds).(1987). *Early Childhood Education: A Developmental Curriculum*. London: Paul Chapman.
- Blenkin, G. & Kelly, A.V. (Eds). (1992). *Assessment In Early Childhood Education*. London: Paul Chapman .
- Bruce, T. (1987). *Early Childhood Education*. London: Hodder & Stoughton .
- Carr, M. (1989). Assessment of what for whom. *Proceedings of Invitational Seminar: Directions for Assessment in Early Childhood*. pp14-30. Wellington: NZCER .
- Carr, M. (1990). To inform their discretion: Assessment in early childhood. In Begg, A. et al (eds) *SAME papers 1990*, pp228-242 Auckland: Longman Paul, 1991.
- Carr, M. (1992a). *Preschool Background Paper. Early Childhood Curriculum Project*. Hamilton: Waikato University.
- Carr, M. (1992b). *Early childhood and school models of curriculum: matches and mismatches. Preschool Position Paper 1. Early Childhood Curriculum Project..* Hamilton: Waikato University.
- Carr, M. (1992c). *Evaluation and behavioural objectives. Preschool Position Paper 2. Early Childhood Curriculum Project*. Hamilton: Waikato University.
- Carr, M. (1992d). *Special characteristics of the preschooler. Preschool Position Paper 3. Early Childhood Curriculum Project*. Hamilton: Waikato University.
- Carr, M. & Dowson, G.(1995). *Te Whaariki: Inclusion and assessment*. Paper presented at the Sixth Early Childhood Convention, Auckland: September 5-8.
- Carr, M. & May, H. (1991). National curriculum guidelines for early childhood in Aotearoa/ New Zealand. In *Proceedings of the Fifth Early Childhood Convention, Dunedin*. pp 212-235
- Carr, M. & May, H. (1993a). The role of government in early childhood curriculum. In V. Podmore (Ed.). (1993). *What is Government's Role in Early Childhood Education? Papers presented at the 1993 NZCER Invitational Seminar, February 23*. pp 42-50. Wellington: NZCER
- Carr, M. & May, H. (1993b). *Te Whaariki Curriculum Papers, Early Childhood Curriculum Project*. Hamilton: Waikato University.

- Carr, M. & May, H. (1994). *Monitoring early childhood experience*. Hamilton:University of Waikato.
- Carr, M.(n.d.) *Some thoughts about linking Parents As First Teachers with the national early childhood curriculum*. Hamilton: Waikato University. Unpublished manuscript .
- Dadds, M.(1993). The Feeling of Thinking in Professional Self-study. *Educational Action Research*, vol 1(2) 287-303.
- DeStefano, L., Maude, S. P.; Crews, S. H. & Mabry, L. (1991). *A statewide search for exemplary practice in early childhood practice in early childhood education*. Paper presented at the Annual Conference of the American Educational Research Association, Chicago, April 3-7. ED 337 246
- Drummond, M.J.; Rouse, D.; & Pugh, G. (1993). *Making Assessment Work: Values and Principles in Assessing Young Children's Learning*. London: NES Arnold/ National Children's Bureau.
- Duckworth, E. (1991). On thinking about teaching. *Educational Leadership*, 48 (6) 30-34.
- Duff, R.E., Brown, M.H. & Van Scoy, I. J. (1995). Reflection and Self-Evaluation: Keys to Professional Development. In *Young Children* May, 81-88, Washington, D.C.:NAEYC.
- Eisner, E.W. (1985a). (2nd ed). *The Educational Imagination: On The Design and Evaluation of School Programmes*. New York: Macmillan.
- Eisner, E.W. (1985b). *The Art Of Educational Evaluation: A Personal View*. Philadelphia: Falmer Press.
- Faragher, J. & MacNaughton, G. (1990). *Working With Young Children: Guidelines For Good Practice*. Victoria:TAFE Publications
- Farquhar, S-E. (1989). Defining and Assessing Quality Early Childhood Centres. In *Invitational Seminar:Directions for Assessment in Early Childhood* pp47-64. Wellington: NZCER.
- Farquhar, S-E. (1993). *Breaking new ground in the study of quality*. Paper presented at the N.Z.A.R.E. Conference, Hamilton, 2-4 December.
- Foote, M., Stafford, P. & Cuffaro, H.K. (1992) . Linking Curriculum and Assessment in Preschool and Kindergarten. In C. Genishi (Ed). (1992). *Ways of Assessing Children and Curriculum: Stories of Early Childhood Practice*. Chapter 6. New York: Teachers' College Press.
- Genishi, C (Ed). (1992). *Ways of Assessing Children and Curriculum: Stories of Early Childhood Practice* . New York: Teachers' College Press.
- Gullo, D.F. (1992) *Developmentally appropriate teaching in early childhood: curriculum, implementation, evaluation*. National Education Association Series. Washington D.C. ED 351 096

- Gullo, D.F. (1994). *Understanding Learning and Assessment in Early Childhood Education*. New York: Teachers College Press, Columbia University.
- Gura, P. (Ed.). (1992). *Exploring Learning: Young Children and Block Play*. London: Paul Chapman.
- Gura, P. (1993). Becoming Connoisseurs and Critics: Making Sense of Block Play. In *Education Development and Care*, vol 92:69-81.
- Haggerty, M. & Hubbard, P. (1994). Te Whaariki Trial: Five Wellington centres work on an agreed basis for curriculum in early childhood services in Aotearoa/ New Zealand.
- Haggerty, M (1996). *Using Video to Work with Te Whaariki: the experiences of five early childhood centres: an interim report*. Wellington: Wellington College of Education.
- Harms, T. (1981). Evaluating Settings For Learning. In M.Kaplan-Sanoff, & R. Yablins-Magid (Eds.). *Exploring Early Childhood: Readings in Theory and Practice*. New York: Macmillan.
- Harms, T. & Clifford, R.M.(1993). Studying Educational Settings. In B. Spodek (Ed.). *Handbook of Research on the Education of Young Children*. New York: Macmillan
- Harms, T., Cryer, D. & Clifford, R.M. (1986). Infant/toddler environment rating scale (Draft). An adaptation of the early childhood environment rating scale. Teachers College Press (1980)
- Honig, A. (1994). *Evaluation of early childhood enrichment programmes*. Paper presented at the Annual Conference of N.A.E.Y.C. , Atlanta, Nov. 30-Dec. 3, ED 377 979
- Hoot, J.; Bartkowiak, E; Goupil, M.A.; Mercado, B.; & Panepento, D.(1993). *A study of the relationship between early childhood teacher evaluation instruments and developmentally appropriate practices/code of ethics criteria*. ED 367 477
- Hurst, V. (1991). *Planning For Early Learning: Education in the First Five Years*. London: Paul Chapman.
- Hurst, V. & Lally, M. (1992). Assessment and the Nursery Curriculum. In G. Blenkin & A. V. Kelly (Eds). (1992). *Assessment In Early Childhood Education*. London: Paul Chapman.
- Kamii, C. (1985). Evaluation: it all depends on your theory. *Australian Journal of Early Childhood*, 10 (2) June, pp 3-9.
- Kaplan-Sanoff, M. & Yaablins-Magid, R. (1981) *Exploring Early Childhood: Readings in Theory and Practice*. New York: Macmillan.
- Katz, L. (1994). Perspectives on the quality of early childhood programmes. *Phi Delta Kappan* November, pp 200-212.

- Katz, L. (1995) Address given at Royal Society for the encouragement of the Arts, Manufactures and Commerce, "Start Right " Conference, London. Unpublished manuscript.
- Killion, J. & Todnem, G.R. (1991). A process for personal theory building. *Educational Leadership*, 48 (6) 30-34.
- Koralek, D. et al, (1993). *The What, Why and How of High Quality Early Childhood Education : A Guide for On-site Supervision*, Washington, D.C.:NAEYC.
- Lally, M. (1991). *The Nursery Teacher in Action*. London: Paul Chapman.
- Leavitt, R.L. & Eheart, B.K. (1991). Assessment in early childhood programs. *Young Children*, July, 4- 8.
- McCrea, N.L. (1989). Regulations: a magic pudding or a swiss cheese? *Australian Journal of Early Childhood Education*, 14 (4) 8-15.
- May, H. (1992). The development of national early childhood curriculum guidelines in Aotearoa-New Zealand: a focus on infants and toddlers. Paper presented at the Bank St. College of Education, Infancy Institute, New York, June 24-26. In Carr, M. & May, H. (1993). *Te Whaariki Curriculum Papers* Hamilton: Waikato University.
- May, H. (1993). Te Whaariki: National early childhood curriculum guidelines in New Zealand. In *Final Report of the International Conference Workshop on Preschool education: A Change Towards the Twenty First Century*. March 20-26, pp116-127.
- Meade, A. (1985). *The Children Can Choose: A Study of Early Childhood Programmes in New Zealand*. Wellington: NZCER.
- Meade, A. (1991). Boffins in Early Childhood Services. In *Proceedings of the Fifth Early Childhood Convention, Dunedin, September 8-12*. pp 55 - 67
- Meade, A. (1993). From curriculum goals to competent children: The decade of quality? In *Proceedings of the CECUA National Curriculum Conference, October* pp
- Ministry of Education (1993). *Te Whaariki: Draft Guidelines for Developmentally Appropriate Programmes in Early Childhood*. Wellington: Learning Media.
- Ministry of Education (1996). *Te Whaariki: Early Childhood Curriculum*. Wellington: Learning Media.
- Mooney, A., Munton, T., Rowland, L. & McGurk, H. (1994). *Developing measures to assess quality in day care*. Poster presented at the biennial meeting of the International Society for the Study of Behavioural Development, Amsterdam, June.
- Moss, P. & Pence, A. (Eds). (1994). *Valuing Quality in Early Childhood Services: New Approaches to Defining Quality*. London: Paul Chapman.

- National Association For the Education of Young Children (N.A.E.Y.C.) (1991). Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 Through 8. *Young Children*, March, 21-38.
- National Research Council (1990). *Who Cares For America's Children?* USA: National Research Council.
- Noori, K. (1994). *A constructivist/reflective paradigm: a model for the early childhood program at Tuskegee University, Tuskegee, Alabama*. ED 370 888.
- Phillips, D.A. (1987). *Quality in ChildCare: What Does Research Tell Us?* Washington, D.C: NAEYC.
- Phillips, D.A. & Howes, C. (1987). Indicators of quality child care: Review of research. In *Quality in ChildCare: What Does Research Tell Us?* NAEYC, Washington, D.C.
- Schweinhart, L.J. (1993). Observing young children in action: the key to early childhood assessment. *Young Children*, July, 29-33.
- Slavenas, R. (1993). Evaluation practices of model early childhood programs. *Early Child Development and Care*, 89, 31-44.
- Smith, A. (1989). Appropriate assessment of young children. In *Invitational seminar: Directions in Assessment in Early Childhood*. Wellington: NZCER .
- Spodek, B. (Ed). (1993). *Handbook of Research on the Education of Young Children*. New York: Macmillan
- Stonehouse, A. (1991). Starting with each child: The challenge of programming for under threes. In *Proceedings of the Fifth Early Childhood Convention, Dunedin, September, 8-12*. pp 100-114
- Torrance, H. (1991). Educational Assessment and Educational Standards: towards and alternative view of quality. In P. Brown & H. Lauder (Eds). (1991). *Education for Economic Survival* . London: Routledge
- Van der Eyken, W. (1992). *Introducing Evaluation: A Practical Introduction to Evaluation in Early Childhood Education Projects*. The Hague:Bernard van Leer Foundation. ED 352 191
- Waters, J.; Frantz, J.F.; Rottmayer, M.T. & Genishi, C. (1992) Learning to See the Learning of Preschool Children. In Genishi, C. (Ed). (1992). *Ways of Assessing Children and Curriculum* . New York: Teachers' College Press.
- Wellington, B. (1991). The promise of reflective practice. *Educational Leadership* 48 (6) 4-5.
- Young, M.E. (n.d.) Investing in Young Children. A World Bank discussion paper, World Bank, Washington.
- Zimiles, H. (1993). The adoration of "hard data": a case study of data fetishism in the evaluation of infant day care. *Early Childhood Research Quarterly*, 8 (3) 369-385.