

[Twist, J. (2002). School Entry Assessment: An Evaluation of its Ability to Support Formative Assessment and the Use of Teacher Judgement. *New Zealand Annual Review of Education*, 11, 215-227]

## School Entry Assessment: An Evaluation of its Ability to Support Formative Assessment and the Use of Teacher Judgement

JULIET TWIST

### **Abstract:**

*This article focuses on the Tell Me (oral language component) of School Entry Assessment (SEA). It finds that although SEA-related documents make a concerted effort to support both formative assessment and the use of teacher judgement in assessment, teachers may not be making formative use of the data obtained, and may be reluctant to use judgement when scoring this component. Suggestions for an improved programme of professional development are discussed.*

In 1997, nationally standardized assessment procedures for use with new entrant children were introduced into New Zealand schools. Entitled *School Entry Assessment* (SEA) in English, and *Aro Matawai Urunga-a-Kura* (AKA) in Maori, the package has three purposes: to provide teachers with information about each child's skills and understanding in selected aspects of numeracy, oral language, and emergent literacy; to help schools decide how best to support new entrants and evaluate their programmes according to identified needs; and to build up a database which will enable the identification of national patterns in numeracy, oral language, and emergent literacy on school entry in order to inform policy development and the allocation of resources (Ministry of Education, 1997, 2001; Gilmore, 1999). The three assessment components of SEA are: *Checkout* (numeracy); *Tell Me* (oral language); and *Concepts About Print* (literacy). Assessment is carried

216 *Juliet Twist*

out by the classroom teacher when the child has been at school between 4-8 weeks.

This article will not evaluate the research underpinning the resource, but will rather focus on an evaluation of the original support manual, *School Entry Assessment: Guide For Teachers* (Ministry of Education, 1997), and the latest progress report, *School Entry Assessment, June 1997–December 2000* (Ministry of Education, 2001). These publications provide teachers with their primary source of information on the philosophy behind the development of SEA, and its intended use. Only the English language package will be considered here, since issues at the heart of experience with the Maori medium resource (concerning its cultural validity and the unique – and sometimes problematic – relationship between teachers in Maori medium settings and the Ministry of Education) cannot easily be discussed alongside the experiences of those working in mainstream settings. However, a recent report to the Ministry of Education on Maori Immersion and Kura Kaupapa Maori teachers' perceptions and use of AKA (Bishop, Berryman & Richardson, 2001) does appear to show similarities with the mainstream experience regarding professional development, and is therefore discussed in both sections of this article.

This evaluation attempts to discover the extent to which SEA (as presented by the 1997 and 2001 Ministry publications) supports assessment for learning, and the extent to which it supports the use of teacher judgement in assessment. The term "assessment for learning", as used here, refers to assessment which is formative in nature. For assessment to function formatively, data must be used to adjust teaching in order to meet learning needs – hence the term. Studies on the efficacy of formative assessment show overwhelmingly, "that innovations which include strengthening the practice of formative assessment produce significant, and often substantial, learning gains" (Black & William, 1998, p. 3). As a consequence, it would appear appropriate to evaluate SEA (a major policy initiative) in terms of its support for this assessment form. This evaluation is also concerned with the extent to which the resource supports the use of teacher judgement. The term "teacher judgement" as used here refers to the involvement of teachers in the analysis or judgement of assessment data in order to discern the next learning step for a child. Willis (1994) claims that the first-hand involvement of teachers in the analysis of data can not only increase their understanding of assessment, but can also contribute significantly to their understanding of the curriculum as a whole.

Because of its potential to enhance professional development, an evaluation of SEA in terms of its ability to support the use of teacher judgement is the focus of the second section.

### Section One: Assessment for Learning

In order for assessment to be formative, it must be primarily focused on the learning needs of the child (Assessment Reform Group, 1999). SEA is not intended for use as a means of making comparisons between children, or between schools; its expressed intention is to provide data that will assist teaching and learning. In *School Entry Assessment: Guide For Teachers* (Ministry of Education, 1997, p. 4), it is stated that “the main purpose of school-based assessment is to improve students’ learning and the quality of learning programmes.” And in the latest report on the resource put out by the Ministry of Education, *School Entry Assessment, June 1997-December 2000*, it is stated that:

Assessment is a vital tool in establishing successful learning. It allows a teacher to identify the skills and understandings children bring into the classroom, and to determine that children are continuing to learn. Assessment also provides opportunities to observe a child and identify the strategies they use, their individual learning styles and learning dispositions. (Ministry of Education, 2001, p. 2)

Nowhere in the introduction does attention veer away from the individual child’s learning needs; it remains resolutely focussed on the importance of accurately identifying the past experiences, skills, understandings, and learning disposition of the child, in order to allow teachers to plan the most effective programme for them. The work of teachers involves continuously observing and assessing progress “to ensure the child is learning and to provide a basis for future learning” (p. 3). The document also provides extensive examples of how schools are using the resource for formative purposes (pp. 14, 22, 30-32).

The publication is divided into four main sections. The first is concerned with learning about the child (not just the academic skills and understandings they have, but also their skills in social interaction, risk-taking, and their learning disposition); the second with how teachers can use the information in order to address learning needs; the third with how the school might use assessment data to assist planning; and the fourth with how all the evidence can be combined in order to form a national picture of the learning needs of young children. All four sections emphasize principles consistent with those on which formative

assessment is based – they all implicitly support the work of Torrance (1993) who argues that formative assessment is designed to assist in the learning process by providing information on pupils’ strengths and weaknesses, and that it must not only identify what pupils have achieved but also what they can potentially achieve given the support of, and interaction with, their teacher.

And yet, according to the Education Review Office (ERO, 1999, p. 13), only 30 percent of schools reviewed in late 1998 used the literacy and oral language components of SEA, and “only a small proportion” were using the numeracy sub-test. While this could, in part, be related to problems concerning the manageability of the resource, with concerns focussed both on the amount of time needed to conduct each assessment component, and difficulties related to administering the package while simultaneously occupying perhaps 25 other children (Dixon & Williams, 2000; Croft, Stratford & Mapa, 2000) – it is not the purpose of this evaluation to address issues of manageability. Those issues need to be addressed, but are not the focus of this article. The intention here is rather to look at the apparent under-use of SEA, and to ask if there is any link between its under-use and the ability of SEA to support formative assessment. That is, the intention is to discover if large numbers of teachers have rejected the resource on the grounds that it does not provide data that can be used formatively, data that can be used to adjust teaching in order to meet learning needs.

Three main pieces of research provide data relating to teachers’ reactions to the (English) SEA package (Goldring, 1999; Williams & Dixon, 1998; Dixon & Williams, 2000; Croft, Strafford, & Mapa, 2000). It should be noted here that the 2001 progress report from the Ministry of Education is not included in the discussion below since, although it contains numerous quotes from teachers discussing their use of the resource, these present a uniformly positive account, suggesting that the report has been designed to encourage teachers to use the resource, rather than to present unbiased data.

For the purposes of this evaluation one element of a study conducted as part of the Research Affiliate Scheme of the University of Canterbury (Goldring, 1999) is particularly relevant. This element involved the investigation of the use of SEA across 56 schools by means of written questionnaires. The comments made by the respondents throw some light on the ERO findings that SEA is underutilized, but fail to support the idea that this under-use is due to the resource not supporting formative assessment.

It is the reactions to the *Tell Me* (oral language) component of the resource that are most revealing, for the purposes of this evaluation. Only six responses related to the strengths of this component, while 30 comments related to its weaknesses. For the most part, dissatisfaction with *Tell Me* centred on the difficulty teachers had with scoring. Of those using the kit, the most commonly cited weakness (offered by 12 respondents) was that the oral language component was difficult to score, with seven respondents using the words “subjective” or “judgmental”. Of those not using the kit, the only comments made about weaknesses concerned the scoring, with two respondents saying it was “subjective” (1999, p. 34). Although not explicitly stated in the research, it can reasonably be inferred that those dissatisfied with the scoring lacked confidence in their ability to judge a child’s competence.

This reluctance to use professional judgement, this dependence upon being given “the answer” strongly suggests that these teachers are unlikely to be able to use SEA as it is intended – as a tool to facilitate the analysis of student understanding in order to diagnose learning need. Although the designers of SEA see the resource as a tool to be used to help teachers accurately identify the skills and understandings new entrant children bring into the classroom, as a tool to help teachers identify how new entrant children make sense of the world, the majority of respondents in this research study seem to be unable to judge, or make sense of, the data provided by the oral *Tell Me* component. It can be suggested at this point, that because these teachers apparently have little faith in their ability to judge a child’s understanding in the oral language context, they will also have difficulty when planning related oral language programmes. The repercussions, however, do not end with oral language, but go further still, since the ability of children to retell a story has a close relationship with their learning and progress in wider areas of language and literacy, such as understanding and creating written stories (Ministry of Education, 1997, p. 22; 2001, p. 4). A teacher unable to judge a child’s competence in story retelling may well be less than effective when planning wider language and literacy programmes.

It thus appears that although the intention of its developers is that SEA should be used to support formative assessment, the inability of a significant number of the teachers in this study to make judgements in the oral language component may prevent this intention from being fully realized. It could, however, be argued that it is the limitations of the resource, rather than those of the teachers, that is at issue. Indeed,

it has been claimed that the 1997 SEA supporting document contains insufficient advice on score interpretation (Croft, Strafford, & Mapa, 2000). If so, when the teachers in the Goldring research raised concerns about making “subjective” or “judgmental” score interpretations, they did so as a result of having been given inadequate advice on which to base their judgements. But the 1997 document offers approximately three full pages on scoring advice for the *Tell Me* component (1997, pp. 26-29), and the 2001 document offers four (2001, pp. 10-13). Although this may not be extensive guidance, there comes a point at which teachers’ inability to make judgements can no longer be attributed to the inadequacies of a supporting document; it perhaps needs to be seen rather in terms of their lack of knowledge of curriculum content.

Williams & Dixon (1998, p.140) have found that a teacher’s ability to use data for formative purposes may be related, in part, to their content knowledge of the particular curriculum area. These researchers interviewed ten teachers using SEA, and found that while half believed the *Tell Me* component gave them useful information about the children’s oral language, “none was able to elaborate in detail how they utilized the information in a formative way.” If, as the researchers suggest, this is partly due to a lack of curriculum content knowledge, it must be asked if it is the function of SEA related documents to raise a teacher’s curriculum content knowledge, or if what is required is a programme of professional development?

It is at this point that research conducted by Bishop, Berryman, and Richardson (2001) into teachers’ perceptions and use of the Maori language version of the resource (AKA) can assist in outlining the design of such a programme. The major recommendation they put forward concerns professional development, and includes one design element in particular which has not, as yet, been part of Ministry-initiated SEA or AKA professional development programmes.

The particular design element highlighted by Bishop et al. is that of the programme *itself* being structured in such a way as to embody the principles of formative assessment, so that trainers and teachers actively practice those principles as they learn how to make the best use of the resource. In accordance with formative assessment principles, teachers and trainers would be required to constantly analyse and evaluate – in a word, “judge” – the efficacy of the learning environment in order to plan the next learning step. On such a programme, the teachers unable to make judgements in the oral language context studied in the Goldring research could be guided towards increasing their curriculum

knowledge within the supportive environment of formative assessment. This would involve trainers and teachers establishing the point at which the learner has prior knowledge; setting mutually agreeable goals; designing authentic contexts for instruction, demonstration, and practice; and taking part in on-going collaborative reflection and feedback.

Unfortunately, just as Bishop and colleagues conclude that AKA-related professional development for teachers has been inadequate, it seems that the story of the mainstream experience of SEA might well read along similar lines. Based on the fact that the respondents in Goldring's research study apparently balked at the expectation that they use their judgement, it can be assumed that their training dealt inadequately (if at all) with this expectation, and that it may also have dealt inadequately with the issue of teacher knowledge of curriculum content.

Although the two Ministry documents evaluated here repeatedly outline the importance of formative assessment, it is apparent that it does not necessarily follow that SEA will be used according to Ministry intentions – especially if programmes of professional development are deficient. Even though the Assessment Reform Group (1999) argues that the focus of assessment now needs to shift away from how teachers *judge* assessment data, and on to how they *use* it, the evidence offered here suggests this approach would be premature. It seems that the use of teacher judgement and the effective use of data are intertwined, that the latter cannot occur without the former, and that to focus primarily on use would be both to overlook the importance of teacher judgement in formative assessment, and to overlook the evidence of teachers' reluctance to use that judgement. Teachers clearly need to learn to appreciate the place of judgement in formative assessment if they are to make the best use of this assessment form, but as is shown above, programmes of professional development – even those initiated by the same body responsible for the development of the resource – do not always adequately deliver. What is needed is a strategy that does not rely on outside initiative, advice, and resourcing, but which can be independently practised by teachers as part of their everyday work. Teacher professional development regarding the use of judgement in formative assessment may, in fact, be most effective when viewed holistically so that it is seen as an integral part of daily practice, and an inherent part of the curriculum as a whole.

## Section Two: Teacher Judgement

The prominent position given to the use of teacher judgement in this evaluation of SEA arises directly as a result of its problematic position in the first section on formative assessment. Ironically, just as the use of teacher judgement appeared as a problem there, in this section it is presented as a strategy for overcoming that problem. It is argued here that exercising judgement – even when one feels a lack of confidence about doing so – has the potential to up-skill teachers to the point where they know enough about their work to be able to confidently judge learning. In a sense here, the problem has become the solution. This argument is based on the work of Willis (1994), where she argues that when teachers are fully involved in all aspects of the assessment process (including using judgement in marking and moderation), they not only learn about assessment, but in addition, also learn about the curriculum in general. So, any assessment package that promotes the use of teacher judgement also makes a significant contribution to teacher professional development. By exercising judgement, through analysing and evaluating, teachers also come to learn about the curriculum as a whole. They will therefore be in a better position to exercise judgement as a result of having initially used it. In the case of the respondents in the Goldring research study, for example, their ability to judge a child's skills and understandings would be enhanced simply through taking a risk and making a judgement. Even if the teachers were not entirely sure about the accuracy of their judgements – and who is? – had they been prepared to enter into the process of judging data they would have – as a result of that process – been up-skilled in the workings of the curriculum to a point where they would have been better equipped to use their judgement. According to this programme of professional development, assessment is not a discrete activity, but is perceived as part of an interconnecting whole. Hargreaves (1989, p. 169), cited by Willis (1994, p. 168), in discussing the mutually reinforcing system formed between teachers, learners, and the curriculum, puts it as follows: "Assessment development, pupil development, teacher development, and curriculum development are ... deeply and inextricably intertwined."

But it is not enough to set up a programme of professional development based upon nothing more than the expectation that teachers use their judgment. It would be unfair – and quite probably counter-productive – to expect teachers to take professional risks

without first ensuring that the environment within which those risks are taken is a supportive one. Such an environment might well be achieved were it to be structured in accordance with the tenets of formative assessment. Just as Bishop et al. argue that programmes of SEA-related professional development initiated outside schools would be more effective if they were based upon the principles of formative assessment, it could also be argued that in-school programmes focusing on the use of teacher judgement would be enhanced if they too were based upon those same principles. If the judging of data began from a point of teacher prior knowledge of curriculum content; if realistic goals were negotiated through dialogue with colleagues; if authentic contexts for instruction, demonstration, and practice were designed; and if ample opportunity for reflection and collaborative feedback were available, a supportive environment (conducive to risk-taking) would become established. Thus a programme of teacher judgement-related professional development designed according to the fundamental principles of formative assessment would set in motion a kind of domino effect. Such a design would in turn enrich teacher understanding of the curriculum, which in turn would make those judgements more accurate and insightful.

While there is evidence that both *School Entry Assessment: Guide for Teachers* and *School Entry Assessment, June 1997-Dec 2000* support the use of teacher judgement, there is a noticeable increase in the importance given to the use of judgement in the more recent document. Although the earlier document repeatedly instructs teachers to “judge” a child’s competence in the sections related to the oral language component (1999, pp. 26-28), the text offers no detail on the significance of this process. In the recent document, however, explicit reference to the use of teacher judgement appears on the first page: “Teachers use the information gathered on entry to school to identify the learning steps and strategies they should offer in the teaching programme.... Such judgements ...” (2001, p. 2) [italics added]. The document portrays the everyday work of teachers as involving the use of judgement, stating teachers will “critically reflect and challenge their thinking and practice”, and that they will “consider” and “evaluate” (pp. 3, 4).

And yet, the weaknesses of the oral language component most commonly cited by respondents to Goldring’s survey were that it was difficult to score, and that it was subjective or judgmental. It would thus appear that because of a reluctance to use judgement these teachers are working to both technocratic and objectivist paradigms – principles that

are patently at odds with the philosophy behind SEA. The technocratic model discussed by Willis (1994) and Cornbleth (1990, p. 31) involves an approach to teaching and learning that:

fosters a reliance on experts and expert knowledge. Teachers (and students) are assumed largely incapable of curriculum planning or enactment in the absence of direction or assistance from curriculum specialists and so-called change agents. The teacher is cast as a manager or passive implementor of expert designs. In effect, technocratic approaches contribute to what has been characterized as teacher de-skilling. A critical approach, in contrast ... assumes and fosters active, reflective, and responsible teacher roles. (1990, p. 31)

In criticizing SEA for its expectation that they should make judgements as they score the *Tell Me* component, the teachers in Goldring’s study also appear to understand pedagogy in objectivist terms, that is, as Cunningham describes objectivism, they “conceive of learning and instruction as phenomena amenable to scientific analysis” (1991, p. 35). The evidence clearly indicates their preference for a resource facilitating the precise measurement of learning. They are in danger of becoming little more than “passive implementors of expert designs” – they will also forgo the opportunity to take part in decision-making that has benefits far beyond the scoring and analysing of a particular score sheet. In using judgement in assessment, a teacher is also necessarily thinking about curriculum development and pedagogy in the broadest sense. In declining to using their judgement, these teachers have in fact lost the opportunity of significant professional development that – if conducted within a supportive environment – could have enabled them to use SEA confidently as its developers intended.

### Conclusion

Perhaps the most striking element of this evaluation is the conspicuous mismatch between the way the significant number of teachers in the Goldring research study used the SEA oral language component, and the way the Ministry intends it to be used. The teachers appear entrenched in technocratic and objectivist models of working, whereas *School Entry Assessment: A Guide For Teachers* and *School Entry Assessment, June 1997 – December 2000* appear to have been produced with a very different use in mind. Indeed the approach of these teachers has far more in common with the ERO report on primary assessment practices (1999, p. 37) than it does with Ministry of Education policy.

The ERO report is critical of teacher skill in recording, analysis, and use of data, saying teachers do not have “sufficient understanding of how the information could be used to guide teaching, [and do not have] sufficient skill in recording and analysis techniques to make it a task they could undertake easily.” But at no stage does the report consider this to be an issue of professional development. The response from the ERO and the teachers in Goldring’s research study is thus, unfortunately, identical – both want to remove the human element from assessment and replace people with a system of assessment characterized by its lack of human input and its pursuit of the objective calculation of learning.

In 1992 the Ministry of Education commissioned a report entitled *A Survey of School Entry Practices: How Schools Gather Information on New Entrants*, which included a recommendation that a working party be set up to:

identify and develop assessment procedures, built on current good practice, which will assist teachers working with new entrants to more accurately *observe* and *interpret* their behaviour and comprehension in a variety of situations covering oral language, reading, and writing ... and mathematics. (Thackery, Syme, & Hendry, 1992, p. 80) [italics added]

*School Entry Assessment: A Guide for Teachers* was the result of that recommendation, and although SEA certainly provides an opportunity for close and detailed *observation* of children as they work, it appears to have been less successful in enabling teachers to better *interpret* the data they obtain. Despite clearly stating the purpose of SEA in terms of formative assessment, and clearly stating the importance of teacher judgement in that process – at least in the more recent (2001) publication – it seems that a reluctance on the part of teachers to use their judgement, and inadequate programmes of professional development, have resulted in practices in schools falling well short of Ministry intentions.

The Goldring study presents a substantial number of teachers as reluctant to use their judgement, and this can only have negative implications for their professional development, their job satisfaction, and most of all, for their students, since teachers who are reluctant to use judgement are unlikely to be able to encourage their students to do so. But if future professional development programmes focus on the use of teacher judgement, in the understanding that the act of making a judgement will require teachers to think more deeply about the

curriculum, with the result that they will consequently be better able to make more informed judgements, *and* if those programmes operate within the supportive framework of formative assessment, the potential of *School Entry Assessment* to improve teaching and learning may come to be fully realized.

## References

- Assessment Reform Group. (1999). *Assessment for learning: Beyond the black box*. Cambridge: Cambridge University Press.
- Black, P., & Wiliam, D. (1998). *Inside the black box: Raising standards through classroom assessment*. London: King’s College London.
- Bishop, R., Berryman, M., & Richardson, C. (2001). *Teachers’ perceptions and use of Aro Matawai Urunga-a-Kura (AKA): Suggestions for improvements*. Wellington: Ministry of Education.
- Cornbleth, C. (1990). *Curriculum in context*. London: Falmer Press.
- Croft, C., Strafford, E., & Mapa, L. (2001). *Stocktake/Evaluation of existing diagnostic tools in literacy and numeracy, in English: A report to the Ministry of Education*. Wellington: New Zealand Council for Educational Research.
- Cunningham, D. (1991). Assessing constructions and constructing assessments: A dialogue. *Educational Technology*, 31(5), 13-17.
- Education Review Office. (1999, Winter). *Student assessment: Practices in primary schools*. Evaluation Reports Number Two. Wellington: ERO.
- Dixon, H., & Williams, R. (2000). School entry assessment: Implementation issues. *set: Research Information for Teachers*, No. 1, 32-36.
- Gilmore, A. (1999). *School entry assessment: The first national picture – July 1997-May 1998*. Wellington: Learning Media.
- Goldring, M. (1999). *School entry assessment: Practices in schools*. Christchurch: University of Canterbury, Education Department.
- Ministry of Education. (1997). *School entry assessment: Guide for teachers*. Wellington: Learning Media.
- Ministry of Education. (2001). *School Entry Assessment: June 1997-December 2000*. Wellington: Ministry of Education.
- Thackery, S., Syme, K., & Hendry, D. (1992). *A survey of school entry practices: How schools gather information on new entrants*. Wellington: Learning Media.

- Torrance, H. (1993). Formative assessment: Some theoretical problems and empirical questions. *Cambridge Journal of Education*, 23(3), 333-343.
- Williams, R., & Dixon, H. (1998). School entry assessment: Praiseworthy or perilous? The tensions between formative and summative purposes. *SAMEpapers*, pp. 135-148.
- Willis, D. (1994). School-based assessment: Underlying ideologies and their implications for teachers and learners. *New Zealand Journal of Educational Studies*, 29(2), 161-174.

**The author**

Juliet Twist has recently completed a Master of Education at Victoria University, and this article is based on an assignment for EDUC547, *Assessment and Evaluation in Action*. Juliet has taught full-time for eight years, including five as Assistant Principal at Petone Central School. Areas of particular interest include literacy, assessment, and transition to school. She is currently working as a relieving teacher in the Wellington area.