

## **Developing a framework for the assessment of musical learning: Resolving the dilemma of the “parts” and the “whole”**

Robert H. Cantwell & Neryl Jeanneret  
School of Education  
University of Newcastle

Recent debate within the domain of music education has focussed on issues of discriminating between higher and lower quality learning outcomes. Much of this debate has centred on the language of music education, particularly in giving both substantive and psychometric meaning to terms as diverse as 'the craft of music', 'musical skills', 'originality', 'musically convincing', and 'convincing development of ideas'. Moreover, in the search for standardisation in music assessment, much of what is conventionally described in assessment criteria reduces musical assessment to quantifiable competencies often not indicative of the higher-order musical thinking underlying the production of these competencies. That is, assessment often fails to resolve the dilemma of the 'parts' and the 'whole'. In this paper we propose an assessment framework based upon a synthesis of current text processing theory with Biggs and Collis= (1982) SOLO Taxonomy. We propose that musical assessment should primarily be sensitive to the quality and structure of music thinking. We argue that musical learning, like other domains of learning, can be analysed for evidence of structural quality and coherence, and that such evaluations may provide viable diagnostic as well as summative information about musical outcomes.

Our purposes in this paper are broadly twofold: to highlight ambiguities and uncertainties associated with current practices in the assessment of creative arts, with particular reference to assessment processes in music; and to propose a framework for such assessment that may give greater clarity and certainty in describing musical performance and outcomes. It is not our intention at this time to propose an assessment schedule for an examination of music - such a schedule would require much empirical validation. What we do aim to achieve, however, is a theoretically defensible framework through which such a schedule may be developed. We argue that such a framework requires some degree of reconceptualisation of what is focussed on in musical assessment, how such foci may reflect the components and their relationships in describing 'good' musical performance, and how, in taxonomic form, such reconceptualized assessment may reliably indicate both quantitative (competencies) and qualitative (expressive) elements of musical outcomes.

We argue that good assessment will achieve three objectives: it will reflect and be informative of the quality of learning/performance outcome, it will reflect and be informative of the processes yielding that outcome, and will reflect and be informative of the cognitive and metacognitive knowledge driving those processes. That is to say, good assessment processes must be descriptive of the quality of learning/ performance and, concomitantly, of those processes and knowledge implicitly indicative of the quality of learning/performance outcome (see Figure 1). Moreover, in domains such as the creative and performing arts (and for that matter, most other areas of learning) comparisons across different performance media and different genre are also required. How does one compare art music composition with popular genre? What similarities of quality of guitar versus woodwind performance may be described and compared? Are there common elements of 'good' performance and 'good' composition that allow for both

quantitative and qualitative comparisons to be drawn? How can common metrics be established that retain both validity and reliability of measurement, yet also in taxonomic form allow for the substantive elements of the performance/composition to be described along with the structural relationships holding these elements together. To a large extent the assessment issue for music, no less than for other creative fields, is one of resolving the dilemma between the ‘parts’ and the ‘whole’ in a form which neither sees assessment as a simple accounting exercise quantifying technical competencies nor as a well meaning exercise in rewarding the vacuous meanderings of an otherwise technically incompetent performance.

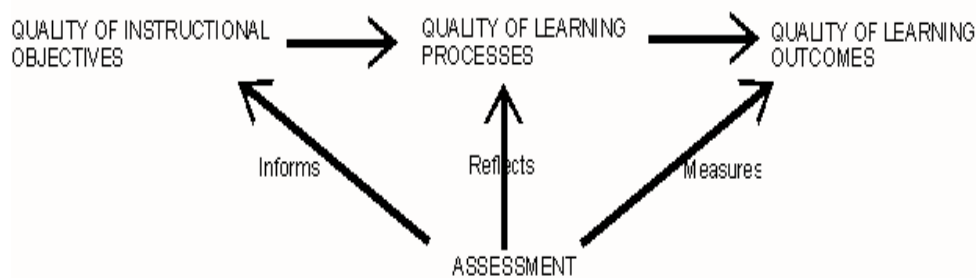


Figure 1: The role of assessment in an integrated curriculum

To a large extent this dilemma has been recognised in the music field. What is less apparent, however, is the manner in which music assessment has been able to deal with the issue of linking the parts to the whole. In a study investigating the use of criteria-specific rating scales for assessing woodwind performance, Saunders and Holahan (1997) constructed a series of measures of technical competencies relating to woodwind performance. These were applied to performance of known score, scales, and unknown score. Additionally, Saunders and Holahan (1997) included a musical interpretation scale for performance on the unknown and known scores (ranging from *The student demonstrates the highest level of musicality including well shaped phrases and dynamics* to *The student demonstrates a lack of musical understanding*).

Each performance was independently adjudicated. Their data suggested high levels of consistency in the scoring of technical competence across all performances ( $r = .92$  across instruments). However, the correlations between the technical competencies and the interpretation scores were only in the moderate range - between .5 and .6 for the known score, and between .4 and .5 for the unknown score. In other words, while the criteria-specific measures appeared to reliably describe technical competencies, the shared variance between these competencies and the interpretation measure suggested that a significant portion of the quality of performance lay in factors additional to the technical competence of the musician. That is to say, while technical competencies may be considered necessary for higher-level musicality, they are not, on the basis of Saunders and Holahan's data, sufficient to explain the quality of performance.

The point of the discussion of Saunders and Holahan's (1997) paper is not to be critical of the adjudication of either the competency ratings or the interpretation ratings. Rather, we are interested in the assessment implications of the finding that the reported levels of technical competence bear only a moderate relationship with the reported levels of musicality. Two issues seem to us to be important here. First, the need to quantify technical competencies in music has not necessarily informed to any significant degree the qualitative aspects of musical performance. Second, while music assessment has recognised the distinction between technical competency and higher order musical meaning, it has not been able to provide a mechanism by which the structural relationships (and indeed predictive relationships) between these may be easily described. We argue that in part this is a reflection of the theoretical assumptions underlying music assessment, and in part is an artefact of the language of music assessment. One approach to the analysis of learning/performance outcome has been the SOLO Taxonomy

developed by Biggs and Collis (1982; 1989). SOLO (an acronym for the *Structure of the Observed Learning Outcome*) is based (in part) on the principle that analysing the structural complexity of a response/ performance will give insight into the quality of the learning outcome or performance. The emphasis placed by Biggs and Collis on qualitative rather than quantitative criteria (such as, for example, the expressive use of technical competencies in performance rather than their mere presence) gives precedence in assessment process to the manner in which the constituent elements of learning and performance are structurally linked. SOLO describes outcomes in terms of a five level hierarchy (see Figure 2) based upon an interaction between the expected modality of performance and the structural complexity of that performance. Both the *prestructural* responses and *extended abstract* levels represent responses outside of the expected modality - in the case of the prestructural response, a failure to address the terms and expectations of the task, and in the case of the extended abstract response, an outcome of such structural complexity as to go beyond the expectations of the task. Examination of the criteria used by Saunders and Holahan (1997) to rate the musicality of the woodwind performance, for example, as well as the criteria established by the New South Wales (NSW) Board of Studies<sup>1</sup> (1999) for the assessment of senior school composition, both reflect SOLO consistent indicators of quality (see Table 1 for examples and DeTurk, 1989; Podgorski, 1989; Cantwell & Millard, 1994 for earlier use of SOLO in musical assessment). However, given our previous contention that good assessment must not only inform of the quality of outcome, but also inform of the processes yielding that outcome and of the cognitive and metacognitive knowledge driving those process, we argue that SOLO analysis in isolation provides only a partial framework for the

---

<sup>1</sup> The NSW Board of Studies provides secondary school exit examinations across a broad range of subjects. Performance in these subjects enables matriculation to university.

<b>Extended Abstract</b>	Taking the task response beyond the limited parameters of the task
<b>Relational</b>	Finding a point of unity relating the key aspects of the data into a meaningful whole
<b>Multistructural</b>	Finding key components of the task and expressing each in an integrated and coherent form
<b>Unistructural</b>	Finding relevant elements of the task and incorporating them in an unintegrated and less coherent form
<b>Prestructural</b>	Unable to address the task in an appropriate or relevant way

**Figure 2:** Biggs and Collis' (1982) SOLO Taxonomy

development of meaningful assessment. While SOLO does indicate the structural quality of the response, and while it also indicates the logic and complexity of the structural relationships between elements within the response, it does nonetheless remain an outcome measure. Understanding of the processes yielding this structure, and of the constituent knowledge reflecting that structure remains largely unspecified. In the rating of both the quality of apparent musicality in Saunders and Holahan's (1997) assessment of woodwind performance, and in the NSW Board of Studies (1999) descriptors of stylistic quality in composition, the conceptual structures underlying the descriptors remain either unclear or unspecified. In neither instance is there objective clarity of what it means to reflect higher-level musicality or higher-level stylistic form. - rather it is the relative presence of these that is emphasised.

Our reconceptualisation of the focus of assessment in music, then, involves a synthesis of SOLO criteria as a measure of global quality with a framework derived from text processing theory. We

Table 1:  
Equivalence of representative assessment criteria in music with Focal levels and SOLO

Focal Level (intention and process)	Evaluation of Woodwind Performance: Musical interpretation. (Saunders & Holahan, 1997)	Composition marking guidelines: Stylistic judgement (Board of Studies, NSW, 1999)	SOLO Equivalent (Biggs & Collis, 1982) (outcome)
<i>Integrative</i>	<p>The highest level of musicality including well shaped phrases and dynamics</p> <p>A high level of musicality but has some phrases or dynamics that are not consistent with the overall level of musical expression</p>	<p>A musically convincing work displaying a high standard of compositional skill and musicality in a style representative of the topic</p> <p>An effective work displaying a reasonable level of compositional skill in a style representative of the topic</p>	<p><i>Extended abstract</i></p> <p><i>Relational</i></p>
<i>Categorical</i>	<p>A moderate level of musicality and musical understanding</p>	<p>A competent work demonstrating a reasonable level of compositional skill in a style representative of the topic</p> <p>Less competent work in a style which may not clearly represent the topic</p>	<i>Multistructural</i>
<i>Reproductive</i>	<p>Only a limited amount of musicality and musical understanding</p> <p>A lack of musical understanding</p>	<p>Demonstrating limited understanding of stylistic features of the topic</p> <p>Work lacks competence, may be incomplete, in a style that does not represent the topic</p> <p>Stylistically incoherent unoriginal work, arrangement only</p>	<p><i>Unistructural</i></p> <p><i>Prestructural</i></p>

argue that this will allow for greater specificity in defining the structure of constituent knowledge underlying the apparent SOLO descriptors. We suggest that such a framework will be informative of both the substantive knowledge underlying different qualities of outcome, and of the processes utilised in the construction of meanings at each of the potential levels of outcome. Cantwell & Millard (1994) have previously alluded to a conceptual relationship between the

quality of thinking in music (as reflected in the planning processes of musicians) and the way in which text is constructed and processed. Current theory in text comprehension/production suggests that learning involves the construction of multi-layered meanings (see Figure 3). That is, information may be processed at three broad levels of analysis: at the level of detail, main ideas and themes. Illustrated in Figure 3 are two crucial features of this conception. First is the hierarchical nature - that is, the transformation of information from detail to main idea and thematic levels is additive: individuals construct meanings that both build on and subsume meanings generated at lower levels. Understanding at the thematic level of analysis, therefore, presumes and requires understandings at the main idea and detailed levels of analysis, just as understanding at the main ideas level of analysis presumes and requires understandings at the

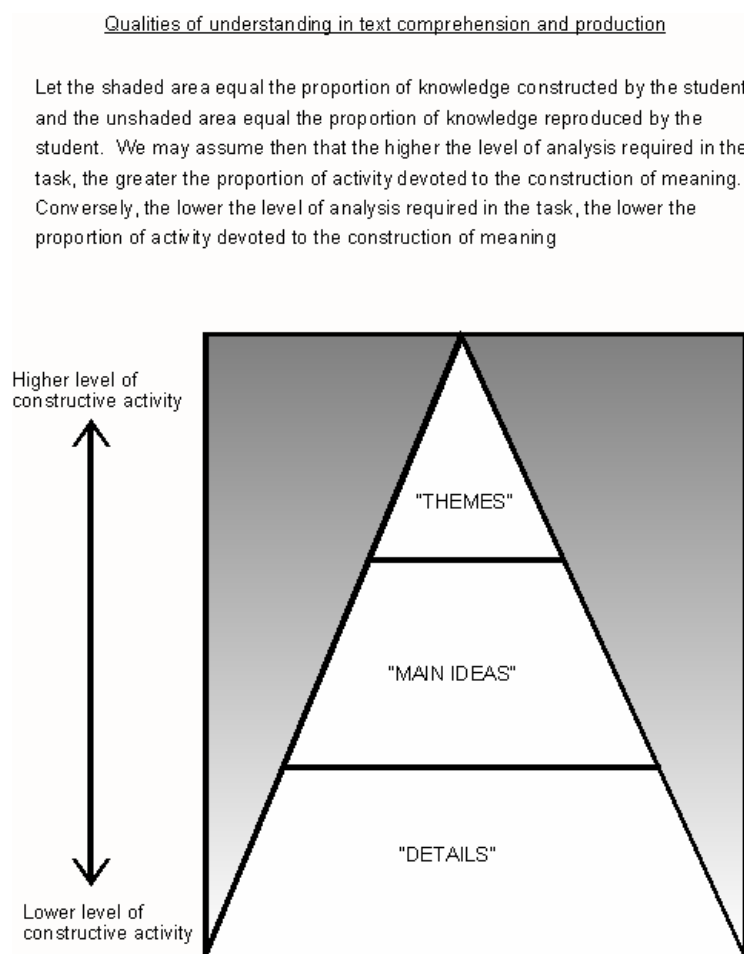


Figure 3: A text processing framework for use in analysing musical products

associated detailed levels of analysis. The second feature is the increasingly active role of the individual learner in constructing meaning at the higher levels of analysis. This is represented in Figure 3 by the grey areas surrounding the triangle. What alters here is the proportion of information which may be taken as given in relation to the proportion of information requiring construction by the individual. The higher the level of analysis, the greater the constructive activity of the individual. Constructing meaning at the thematic level, then, ultimately calls upon a significant amount of prior knowledge and prior understanding as the reference point for making higher order sense of the text or data under consideration. Moreover, as research by Sullivan and Cantwell (1999) has demonstrated, higher order musical processing involves not only significant prior knowledge and understanding about music, but also more general dispositions towards meaning creation that are not necessarily specific to the domain of music itself.

Based on the notion of differential levels of meaning construction in text analysis, we propose that more generic qualities of process and outcome may be inferred. Detail level analysis leaves information largely untransformed, with the consequence that learning outcomes may rarely reflect more than a *reproductive focus*. Main idea level analysis represents a first order transformation of information. It is, however, a *categorical focus* in which information is largely summarised under context dependent and conventional headings. The thematic level of analysis involves a more fundamental transformation of information through the construction of meanings beyond the literal meanings exemplified in the main idea and detail levels of analysis. That is, analysis involves an *integrative focus*, potentially extending meanings to quite high levels of abstraction.

Our assessment framework, then, specifies three generic qualities of process (the reproductive, categorical and integrative foci) reflective of five qualities of outcome (from prestructural to extended abstract), each of which implies different qualities of knowledge use and strategic behaviour. At the same time these focal levels may also be seen as interactive and additive through the construction of coherent relationships within and between knowledge elements constructed at each level of analysis. Outcomes reflective of an integrative focus, for example, would both imply and subsume outcomes generated at the categorical and reproductive levels. The equivalence of the text processing/construction model and the more generic focal framework is illustrated in Figure 4.

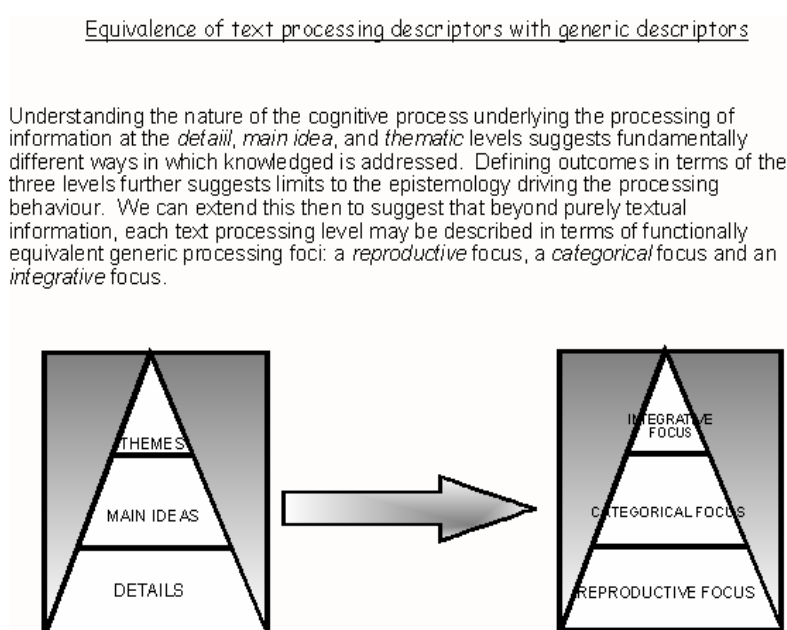


Figure 4: From structural complexity to attentional focus

To what extent does this framework inform assessment in general and assessment in music in particular? One of the deficiencies of current assessment practice identified earlier in this paper was the difficulty in linking the higher order judgements of musical quality with lower level (but

nonetheless important) components of musical processing. In the particular example from the Saunders and Holahan (1997) study the relationship between ‘lower order’ technical competencies and ‘higher order’ interpretative activity was at best moderate. A significant proportion of the variance in interpretative quality was left unexplained by reference to technical performance. This disparity suggested that other factors, in addition to technical competence, might influence the quality of musical interpretation.

How can we explain this disparity by reference to the hierarchical assessment framework?

Interpretation in the sense used by Saunders and Holahan (1997) refers to the degree of ‘musicality’ evident in the performance as judged by the adjudicators. Yet even in their study, the meaning of ‘musicality’ remained largely undefined beyond a sense of what each adjudicator subjectively perceived to be ‘musical’. So in applying the three level taxonomic framework to Saunders and Holahan’s (1997) data (see Figure 5), the inferred relationship between interpretation as indicative of an integrative focus, and the creation of a technically competent whole as would be inferred by higher level technical performance at the categorical level, remained largely unestablished. Whilst technical competence across all the elements identified by Saunders and Holahan (1997) no doubt contribute to the overall musicality of performance, they do not in themselves provide sufficient conditions to explain musicality.

The critical feature of the of the assessment framework is that information constructed at each level of meaning coheres through the establishment of logical links both within and between each focal level. That is to say, what exists at one level is directly and logically derivative of information constructed at lower levels. The integrative level, for example, is derivative of information constructed at the categorical level. The difficulty with the Saunders and Holahan

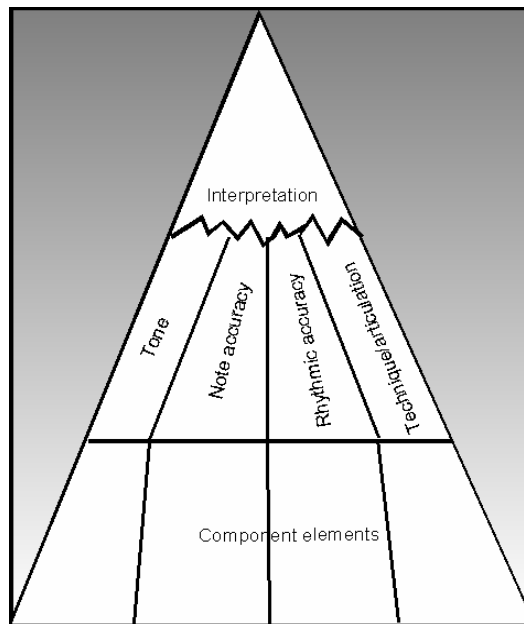


Figure 5: Application of the taxonomic framework to Saunders & Holihan's (1997) data

(1997) data is not that at the categorical level technical competencies are irrelevant, nor that at the integrative level that musicality is not descriptive (at least in SOLO-like terms), but that the empirical relationship between the two is itself not adequate to explain the quality of outcome. In other words, explanation of performance at the integrative level requires consideration of information additional to, and interactive with, the technical competencies underlying musical production.

Underlying the taxonomic model of attentional focus is the assumption that *intention* in learning and performance will typically constrain processing activity to particular levels of meaning. Given the theoretical and empirical link between intention and process described in the student learning literature (eg. Biggs, 1996), shifts in attentional focus by the musician (such as from a reproductive to a categorical focus) represents more than a simple additive process - rather such a shift represents a transformational process driven by qualitatively different conceptions of the

possibilities and purposes of the musical process itself. As illustrated in Figure 2, the shift to a higher level of attentional focus is necessarily accompanied by an increased proportion of attention devoted to constructed as opposed to given information. The operative question therefore becomes one of defining what other elements of musical knowledge beyond the technical contribute at the categorical level to an integrative focus, and what attributes of an integrative focus imply and are derivative of the information constructed at the categorical level.

It is our argument that the coherence between focal levels, particularly between the categorical and integrative levels, involves greater specificity in two ways. First, information utilised at the categorical level must incorporate information beyond the given stimulus field. This may, for example, include such elements as prior strategic knowledge and the process of strategic selection (Cantwell & Moore, 1996; Sullivan & Cantwell, 1999), the quality, structure and accessibility of prior domain knowledge and its links to current task parameters (Cholowski & Chan, 1994; Irvine & Cantwell, 1999; Irvine, Cantwell & Jeanneret, 1999), as well as efficacy judgements (eg. Bandura, 1993) and dispositional biases in the way in which the task is appraised and engaged (Salomon & Globerson 1987; Sullivan & Cantwell, 1999) (see Figure 6).

By implication, then, assessment of the quality of learning must incorporate understandings of what Richard Snow (1989) has referred to as the “cognitive and conative structures in learning@ (p1). As Snow (1989) has also suggested, the utility of this conception in aiding the interpretation of the quality of the learning outcome, both summatively and diagnostically, requires careful attention to the principle of construct validation. We would argue that consideration of the reproductive, categorical and integrative elements framework provides a theoretical rationale for deconstructing quality performance into the underlying predictive components (including technical competencies). It is this capacity to specify the logical

relationships within and between

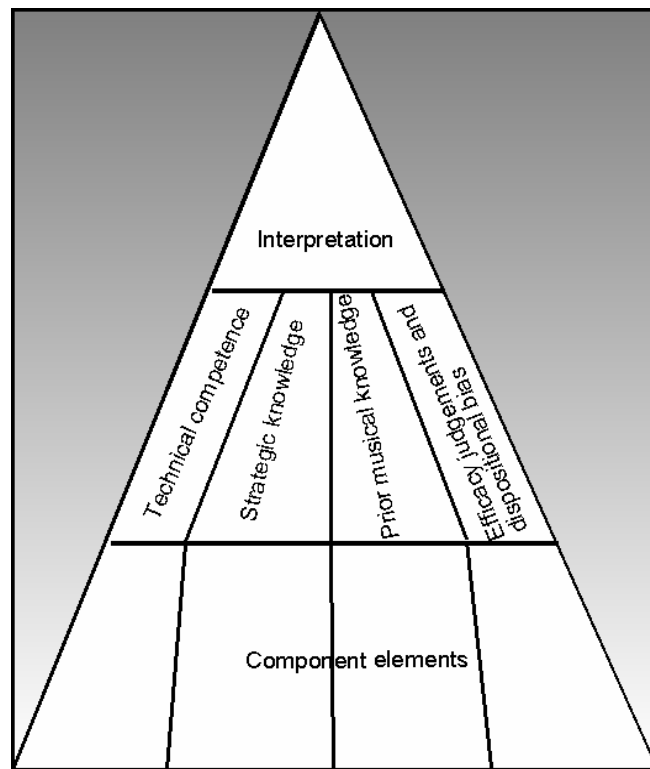


Figure 6: An expanded framework for a taxonomic description of musical quality

the different levels of meaning that allows for the construction of testable assessment hypotheses and through this the construction of defensible assessment schedules.

The second area of specificity implied by our approach is the need to clarify what is meant by descriptors at the integrative level. Descriptors utilised in the assessment of higher order musical outcomes need not only to be transparently indicative of certain qualities of outcomes, but must also be transparently derivative of the categorical information underlying and producing that quality of musical thought. Composition marking guidelines for senior school music in NSW (Board of Studies, 1999) specify outcomes at five possible levels of achievement, ranging from

‘poorest’ (1 B 3 out of a possible 20 marks) to ‘outstanding’ (18 B 20 out of a possible 20 marks) with a normal distribution of marks spread over the other three levels. The assessment rubric specifies a set of performance indicators for each outcome, which in turn act to inform levels of achievement. At the same time, the operationalising of these indicators is often associated with discrepancies between the two marking teams. We suggest that such discrepancies may exist as a result of the structural confusion between indicators and achievement (see Figure 7). With limited space and with confusion between indicators and levels of achievement, statements such as “sustained involvement in the composition process” appear as an ‘outstanding’ level of

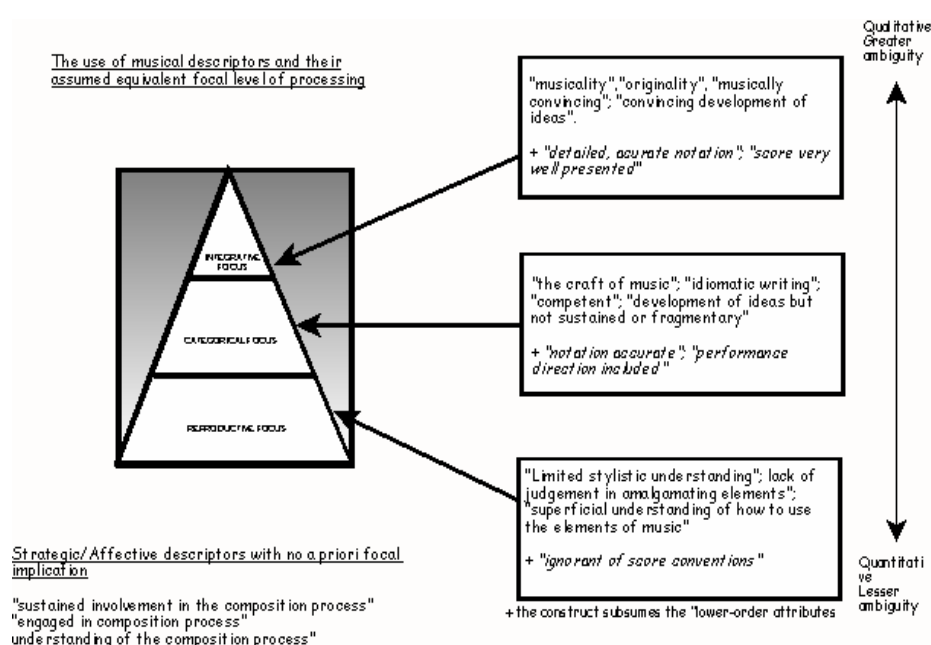


Figure 7: Discriminating between indicators and achievement in assessment.

achievement for the outcome *demonstrated understanding of the composition process*. The levels of quality of outcome, processes that yielded this outcome, and cognitive and metacognitive knowledge employed in this process remain unclear, as do the physical attributes of the work that deem it ‘outstanding’. Moreover, the guidelines rest on the assumption that all the outcomes

are equal and that ‘poor’ achievers will consistently demonstrate the descriptors in relation to each of the outcomes. In reality markers often debate the placement of a work when the candidate is achieving different levels for different outcomes. For example, it is possible for a candidate to produce “stylistically incoherent work” (poor) with a “detailed and accurate score” that indicates ‘outstanding’. We would argue that “notation@ and ‘coherence of presentation” relates to a knowledge of musical conventions which is, in turn, part of the “understanding of the concepts of music”. Scoring issues then become part of a more substantial and comprehensive outcome instead of being given an unjustified, equal weighting.

In summary, we have proposed a framework for the assessment of music that attempts to situate such assessment within the broader rubric of current theoretical work in cognitive and educational psychology. It has not been our intention to provide a schedule for assessment in music, but rather to suggest a framework within which such assessment may be developed. Given the emphasis in current musical assessment on issues of both quality of outcome and of technical competencies, we suggest that it is crucial to effective music assessment to utilise a framework where the two perspectives are seen to be substantively and psychometrically compatible. Our reference to the three focal levels of attention and processing in any musical endeavour allows for both the cognitive and conative dimensions of performance to be identified and contextualised. Moreover, as we illustrate in Figure 8, such a framework allows for the tracking of both higher and lower quality outcomes into their constituent elements without compromising understanding of the relationships between these components and the quality of broader musical outcome (see Table 1). That is, we suggest our framework allows for a first step in resolving the dilemma of the ‘parts’ and the ‘whole’ in the assessment of music.

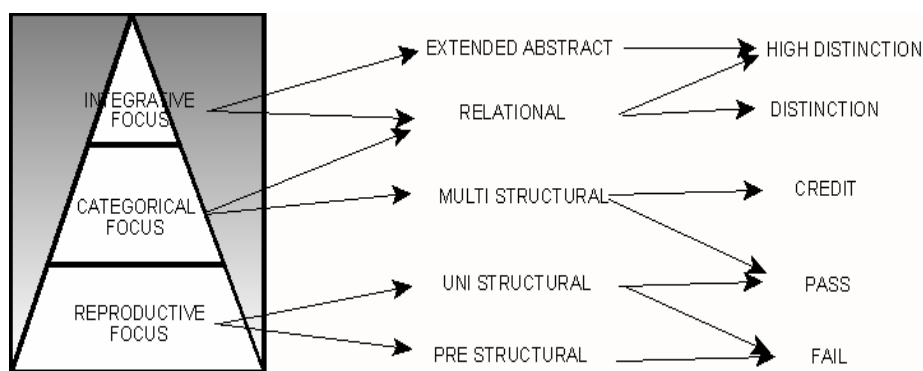


Figure 8: An integrated framework for assessment

## REFERENCES

- Board of Studies, NSW (1999). *1998 HSC Examination Report: Music*. Sydney: Board of Studies.
- Biggs, J. (1996). Assessing learning quality: reconciling institutional, staff and educational demands. *Assessment and Evaluation in Higher Education*. 21, 5-15.
- Biggs, J. & Collis, K. (1982). Towards a model of school-based curriculum and assessment using the SOLO taxonomy. *Australian Journal of Education*, 33, 151-163.
- Biggs, J. & Collis, K. (1982). *Evaluating the Quality of Learning: The SOLO Taxonomy*. New York, Academic Press.
- Cantwell, R. & Moore, P. (1998). Relationships among control beliefs, approaches to learning, and academic performance of final year nurses. *Alberta Journal of Educational Research*. 48, 95-99.
- Cantwell, R. & Millard, Y. (1994). The relationship between approach to learning and learning strategies in learning music. *British Journal of Educational Psychology*. 64, 47-65.
- Cholowski, K. & Chan, L. (1995). Knowledge-driven problem-solving models in nursing education. *Journal of Nursing Education*. 34, 148-54
- DeTurk, M. (1989). Critical and creative musical thinking. In E. Boardman (Ed) *Dimensions of Music Thinking*. Reston, Virginia: MENC. 21 - 32
- Irvine, I., Cantwell, R. & Jeanneret, N. (1999). *Individual differences in processing behaviours of three tertiary students engaged in music composition*. Paper presented at the Joint AARE/NZARE Conference, Melbourne, November.
- Irvine, I. & Cantwell, R. (1999). *Musical composition: Toward a theoretical model*. Paper presented at the ASME Conference, Sydney, July.
- Pogonowski, L. (1989). Metacognition: A dimension of music thinking. In E. Boardman (Ed) *Dimensions of Music Thinking*. Reston, Virginia: MENC. 9 - 19.
- Salomon, G. & Globerson, T. (1987). Skill may not be enough: The role of mindfulness in learning and transfer. *International Journal of Educational Research*. 11, 623-638.
- Saunders, T. Clark & Holahan, J. (1997). Criteria-specific rating scales in the evaluation of high school instrumental performance. *Journal of Research in Music Education*. 45, 259-72.
- Snow, R. (1989). Toward assessment of cognitive and conative structures in learning. *Educational Researcher*. 18(9), 8-14.
- Sullivan, Y. & Cantwell, R. (1999). The planning behaviours of musicians engaging traditional and non-traditional scores. *Psychology of Music*. 27, 245-266