DOCUMENT RESUME

ED 441 839 TM 030 874

AUTHOR Jiang, Ying Hong; Smith, Philip L.
TITLE A Construct-Centered Generalizability Model: Analyzing Underlying Constructs of Cognitively Complex Performance Assessments.
PUB DATE 2000-04-00
NOTE 49p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, LA, April 24-28, 2000).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Cognitive Tests; *Construct Validity; Elementary Education; *Elementary School Students; *Generalizability Theory; *Mathematical Models; *Performance Based Assessment; Reliability; Scoring

ABSTRACT

With a construct-centered reliability analytical approach the reliability analysis should crystallize the multi-traits or constructs that the test specialists developed to measure from student performance and then estimate the degree of fit between the theoretical expectations of test developers and the performance exhibited by students. This study attempts to provide an empirical example of a construct-centered reliability analysis using writing performance assessment data from a large urban school district. Writing assessments from 17,330 students in grades 3, 5, and 8 were used in the analysis. Scoring guides were developed, centering the constructs comprised of dimensions for raters to consider when scoring. Resampling was done by randomly selecting three of four percent of the students from the assessments before a generalizability study with a fully crossed two-facet design was conducted. A large proportion of variance was estimated due to the constructs of rhetorical effectiveness and conventions, as well as the interaction between constructs and raters. Variance due to rater facet was very small. An array of acceptable G coefficients across samples was obtained. It is suggested that high reliability or generalizability is achievable using a construct-centered reliability approach to identify construct relevant variances when there is a high degree of fit between the substantive expectations generated from the test specialists' understanding of the construct as realized by the measurement procedure and observation. (Contains 7 tables, 12 charts, and 10 references.) (SLD)

A Construct-Centered Generalizability Model: Analyzing Underlying Constructs of Cognitively Complex Performance Assessments

Ying Hong Jiang
Long Beach Unified School District

Philip L. Smith
University of Wisconsin-Milwaukee

Paper Presented at the 81st Annual Meeting of the American Educational Research Association

April 24-28, 2000

New Orleans, LA

Abstract

Performance assessment differs from traditional assessment in that it claims to measure multi-traits simultaneously rather than consistently measure single trait. Traditional reliability analyses oversimplify issues of generalizability and stability when applied to cognitively complex performance assessments. Nichols and Sugrue (1998) recommend that reliability analyses incorporate the complex cognitive assumptions so that performance assessments more faithfully represent the multidimensional constructs common to performance assessments. More specifically, with cognitively complex assessments, reliability or generalizability must be reconceptualized to reflect the theoretical expectations of the test developer as well as the complex thinking of the test takers (e.g. problem solving, communication of ideas; reasoning etc.) (Nichols and Smith, 1998). With a construct–centered reliability analytical approach, the reliability analysis should crystallize the multi-traits or constructs that the test specialists developed to measure from student performance, and then estimate the degree of fit between the theoretical expectations from test developers and the performance exhibited by the students.

This study attempted to provide an empirical example of a construct-centered reliability analysis using writing performance assessment data from a large, urban school district. Approximately 31,645 writing assessments taken in English during the spring of 1998 at 6 grade levels. Grade 3, 5, and 8 writing assessments were used in the analysis. The data used for this analysis contains student responses to one writing prompt for each student. The assessments were designed focusing on the constructs underlying different domains tested. Meanwhile, scoring guides were developed centering the constructs comprised of dimensions for raters to consider when scoring. Raters were trained to follow the construct-centered-scoring guide. Resampling was done by randomly selecting 3-4 % of the students involved in the assessments before a generalizability study with a fully crossed two-facet design was conducted. Large proportion of variance was estimated due to the constructs of rhetorical effectiveness and conventions as well as the interaction between constructs and raters. Variance due to rater facet was very small. An array of acceptable G coefficients across samples was obtained. It is suggested that high reliability or generalizability is achievable using construct-centered reliability analytical approach to identify construct relevant variances where there is a high degree of fit between the substantive expectations generated from the test specialist's understanding of the construct-as realized by the measurement procedure-and observation (Nichols and Smith, 1998).

A Construct-Centered Generalizability Model: Analyzing Underlying Constructs of Cognitively Complex Performance Assessments

Ying Hong Jiang, Long Beach Unified School District
Philip L. Smith, University of Wisconsin-Milwaukee

The Methodological Challenge

Generalizability problems in cognitively complex assessments. Traditional reliability analyses oversimplify issues of generalizability and stability when applied to cognitively complex performance assessments. To add to the challenge of estimating reliability or generalizability of cognitively complex assessments, proponents of both performance and portfolio assessments build them to measure multidimensional constructs (Romberg and Wilson, 1995; Wiggins, 1993).

Unfortunately, a large body of research (Koretz, McCafferey, Klein, Bell, & Stecher, 1992; Shavelson, Gao, & Baxter, 1993, etc.) reports low reliability or generalizability indices for performance assessments. The multidimensionality of constructs, which we call "cognitive complexity" is hinted at in traditional generalizabiltiy analyses. Jiang, Smith and Nichols (1997) conducted a meta-analysis of 22 studies using cognitively complex performance tests and found that between-task variance was the major facet, and significantly larger than between-student variance. Task variance, under traditional conceptualizations of reliability, is something to be "factored out" or "equalized" rather than incorporated into generalizability estimates. The problem with task variance "swamping" the variance associated with students creates a "Catch-22" for developers of performance assessments who create measures under different assumptions than those created by random sampling theory.

Construct Centered Generalizability. Recent work in reliability theory proposes that in cognitively complex assessments, the distinction between "reliability" and "validity" becomes blurred. The same conclusion was reported in validity theory as well (Cronbach, 1988, 1989 ; Messick, 1989) with the "unitary" concept of validity. Nichols and Sugrue (1998) recommend that reliability analyses incorporate the complex cognitive assumptions so that performance assessments more faithfully represent the multidimensional constructs common to performance assessments.

More specifically, with cognitively complex assessments, reliability or generalizability must be reconceptualized to reflect the theoretical expectations of the test developer as well as the complex thinking of the test takers (e.g. problem solving, communication of ideas; reasoning etc.) (Nichols and Smith, 1998). The logical implication for practice is that test developers should incorporate both the 1) theoretical expectations/constructs to be tested and 2) multi-trait, complex thinking required of the test takers into their reliability analyses.

How Might Construct Centered Reliability Work?

Overview of the study. Conceptually, a performance assessment score may be viewed as a sample of student performance drawn from a complex universe defined by a combination of all the admissible tasks, occasions, raters, and measurement methods. Generalizability theory refers to each of these dimensions as "facets". The task facet represents the content in a subject-matter domain; the occasion facet includes all possible occasions on which a decision maker would be equally willing to accept a score on the performance assessment. The rater facet includes all possible individuals who are trained to score the performance reliably. Typically, these three

facets are viewed as primary sources of error in a measurement procedure, especially a performance assessment.
Traditional reliability analysis is concerned with the homogeneity of the tasks comprising an assessment. Assessments based on traditional measurement models are intended to measure a single phenomenon, or a 'unitary trait'. The dependability of an assessment is assessed by a reliability or generalizability coefficient, which is a function of the degree of homogeneity of tasks and number of tasks. Within traditional reliability frame, a highly dependable assessment is comprised of a sufficiently large number of homogeneous tasks. The conditions to grant acceptably dependable assessments impose a difficult challenge to performance writing assessments, as it is not only inappropriate but also impossible to administer 'a sufficiently large number of homogeneous writing tasks' to a large number of students.

Performance assessment differs from traditional assessment in that it claims to measures multi-traits simultaneously rather than consistently measuring a single trait. With a construct–centered reliability analytical approach, the reliability analysis should crystallize the multi-traits or constructs that the test specialists developed to measure from student performance, and then estimate the degree of fit between the theoretical expectations from test developers and the performance exhibited by the students. Hence, a good scoring rubric should reflect well the theoretical expectations or the multi-traits or constructs the assessment is developed to measure. Since one assessment is comprised of multi-traits rather than single traits, we should further decompose or unfold tasks into mini-tasks, which are the smaller and more homogeneous dimensions that comprise the multi-traits or constructs for each performance assessment. Since there are mini-tasks nested within the multi-traits, the level of reliability analysis should reflect this hierarchical structure. Further more, when reliability analysis incorporates the theoretical expectations from test specialists, more variance is defined systematically rather than left undefined as residue or error variance, thus, granting higher assessment dependability using traditional methods.

The purpose of this study was to provide an empirical example of a construct-centered reliability analysis using writing performance assessment data from a large, urban school district. Approximately 31,645 writing assessments were taken in English during the spring of 1998 at 6 grade levels. We used grade 3, 5, and 8 writing assessments in the analysis.

Sample and Methods

Sample. The district from which our study data were taken administers writing performance assessments in grades 3, 5, 6, 8, 10 and 11 each spring. The sample data used were from Spring 98. The assessments are given to all students at the grade level and used for program evaluation purposes. Two prompts from the same writing domain are administered at each grade level. Students are randomly assigned to prompts.

Table 1 displays the number of assessments, writing domains assessed in English and grade levels included in the study.

Table 1
Number of Student Writing Assessments in English by Grade Level

Grade Level	Writing Domain Tested	Number of Assessments
3	Autobiographical Incident	5488
5	Observation	6009
8	Speculation about Cause-Effect	5833
TOTAL		17330

The student population in our sample is quite diverse. The district is 54.8% Free and Reduced Price Lunch, 40.7% Hispanic, 20.2% African American, 14% Asian and, 19.4% white. The percent of English Language Learners ranges from 40.7% at the elementary level to 29% at Middle and 22.3% at high school. Table 2 summarizes the demographic characteristics of our sample. Because the entire population was sampled at the grade levels included in our study, the sample demographics represent the district profile quite well, except for the sample only included assessments in English, the proportion of English language learners is lower in the sample than district population (see Table 2, Student Sample Demographics).

-----(Insert Table 2 here)-----

The assessments. Students in grade 3, 5, and 8 took writing assessments that were intended to test specific domains with several simultaneous underlying constructs. Table 3 summarizes the constructs by grade level and writing domain (also see the scoring guide Table 4, 5, and 6 in appendix).

Table 3
Constructs By Writing Domain

Grade	Writing Domain	Rhetorical Effectiveness Construct Measured	Knowledge	Conventions Constructs Measured	
3	Autobiographical incident	Incidence	N/A	Usage	Spelling and Mechanics
		Context			
		Voice and style			
		Significance			
5	Observation	Identification of subject	N/A		
		Context			
		Observational stance			
		Presentation of experience			
8	Speculation about causes and effect	Presentation of situation	Knowledge		
		Logic and relevance of causes			
		Elaboration of argument			

Chart 1 explains the relationship between the domain tested for grade 3, 5 and 8 and the constructs underlying the domains as reflected in the scoring guide.

--- (Insert Chart 1 here)---

Scoring. Teachers were released district wide for one day in late May to score student assessments and were asked at the end of the scoring day to summarize the learning issues they encountered in the papers. Teachers were assigned to rate papers at their own grade level and trained prior to rating on two scoring guides, one for Rhetorical Effectiveness and one for Conventions (see Appendix for copies of the scoring guides). The rater training consisted of one to two hours of guided rating and discussion managed by a district-trained and experienced Table Leader. The Table Leader discussed the prompt and scoring guide, had teachers write to the prompt (depending upon the experience of the group being trained), and facilitated the rating "anchor papers" (previously selected and scored papers that represent the score points in the Rhetorical Effectiveness and Conventions scoring guides). Each student received one score on

Rhetorical Effectiveness and one on Conventions. Scores were from 1 to 6, with 6 being the highest. Once the Table Leader was satisfied that the teachers reached consensus about the scoring guides, rating began. Papers from each of the districts 62 elementary and 21 middle schools were randomly assigned to tables of teachers. Teachers gave each paper a first read. During the first reading, Table Leaders "read behind" teachers and if a teacher was straying from the scoring guides, she was pulled aside for re-training. Papers were then collected and randomly assigned to another set of teacher raters for a second reading. Table Leaders checked the second reads. When scores were more than two points apart, the Table Leader gave the paper a third reading and that "expert" score became the score reported to the student.

Teachers assigned one score for Rhetorical Effectiveness and one for Conventions. In determining scores, teachers considered grade-level specific constructs defined by numerous dimensions. In the third grade writing domain of autobiographical incident, for example, rhetorical effectiveness was divided into four constructs: incidence, context, voice and style, and significance, and the conventions was divided into two constructs: usage, spelling and mechanics. The scoring guide directed teachers to evaluate incidence along 34 dimensions; context by 11 dimensions; voice and style by 18 dimensions and significance by 16 dimensions.

1) Grade 3 data

For grade 3, the domain tested is AUTOBIOGRAPHICAL INCIDENT. The two prompts used were **Form A**: *A Time I Learned How To Do Something*, and **Form B**: *A Time I Didn't Give Up*. The four underlying constructs to measure AUTOBIOGRAPHICAL INCIDENT are ***incidence, context, voice and style,*** and ***significance***. According to the scoring guide, there are 34 dimensions to measure the construct of incidence.

-----(Insert Chart 2 here)----

From the chart, we can assume that a score of 6 implying that a student has met the 11 dimensions (IQ1 to IQ11) developed according to the scoring guide. Likewise, for a score of 5, we can assume that a student has met the 4 dimensions (IQ12 to IQ15), and etc. Chart 3, 4, and 5 shows the dimensions indicating the construct of context, voice and style, and significance separately.

-----(Insert Chart 3, 4, and 5 here)-----

2) Grade 5 data

The domain tested for grade 5 is OBSERVATION. The two prompts used were: **Form A**: *A Rainy Day,* and **Form B**: *Watching an Animal*. The four underlying constructs to measure OBSERVATION are ***identification of subject, context, observational stance,*** and ***presentation of the experience***. According to the scoring guide, there are 5 dimensions to measure the construct of identification of subject, 6 dimensions to measure context, 7 dimensions to measure the construct of observational stance, and 18 dimensions to measure presentation of the experience. For each score category from 1 to 6, there are corresponding dimensions developed to reflect if a student has met those dimensions. Chart 6, 7, 8 and 9 shows the dimensions indicating the 4 constructs

-----(Insert chart 6, 7, and 8 here)-----

3) Grade 8 data

SPECULATION ABOUT CAUSES AND EFFECT is the domain tested for grade 8. The two prompts used were **Form A**: *Lewis and Clark,* and **Form B**: *Overland Trails to the West.* The three constructs underlying to measure the domain were ***presentation of the situation, logic and relevance of causes,*** and ***elaboration of argument.*** Chart 10, 11, and 12 show there are 36 dimensions are developed to measure the construct of presentation of the situation, 40 dimensions to measure logic and relevance of causes, and 31 dimensions to measure elaboration of argument.

-----(Insert chart 10, 11, and 12 here)-----

Analytical Procedure

Re-sampling. Each grade, approximately 5,000 to 6,000 students were involved in the assessments. We randomly sampled 3-4 % assessments from the entire database three times with replacement before we obtain all the parameter estimates for analysis.

Generalizability analysis. We employed generalizability theory (Cronbach, Cleser, Nanda, & Rajaratnam, 1972) to analyze the dependability of the assessments. Dependability usually refers to the accuracy of generalizing from a person's observed score on a test to the average score that person would have received under all the possible conditions that the test user would be equally willing to accept. The strength of G theory is that multiple sources of error can be estimated separately in a single analysis. We designed G study with two facets to focus on the variance components due to constructs comprising writing domains based on observed scores on rhetorical effectiveness and conventions, and the variance components due to constructs and rater interactions.

For grade 3, 5, and 8, we designed a (2x2) two facets fully crossed design, with two constructs (rhetorical effectiveness and conventions) and two raters. For grade 8, although knowledge was one of the constructs underlying the domain tested, we consider it a different construct from rhetorical effectiveness and conventions since it measures mastery of historical knowledge.

Table 4 shows the generalizability analysis design for grade 3 and 5, and 8.

Table 4: G Study Two Facets Fully Crossed Design for Grade 3, 5, and 8:

Student	*Rater*	*Rhetorical*	*Convention*
	1	*score*	*score*
	2	*score*	*score*

Results and Evidence

The estimates of variance components from separate sources of variation are shown in Tables 6, 7 and 8. For grade 3, the variance due to construct is 10%, and the variance due to the interaction between constructs and raters is around 7%. For grade 5, the variance due to construct is above 7% and the variance due to the interaction between constructs and raters amounts to around 20%. For grade 8, even larger proportion of variance due to constructs is observed to be around 48.49%. The variance due to the interaction between constructs and raters is around 2.18%. It is noticeable that the variance due to rater is very small.

G coefficients for each grade were derived to estimate the proportion of expected-score variance that is the universal score variance. For grade 3, the range of the G coefficients for samples A, B, and C were from .747 to.788; for grade 5, the range of the G coefficients for samples A, B, and C were from .677 to .763; for grade 8, the range of the G coefficients for samples A, B, and C were from .809 to .821. Except for the G coefficient for grade 5 sample B was a little lower, the rest of the G coefficients fall into acceptable range.

Table 6: Estimates of Variance Components for G Study Two Facets Fully Crossed Design

(Grade 3)

Source of Variation	Estimated Variance Components				
	Sample A (N= 228)	Sample B (N=213)	Sample C (N=213)	Overall Average	
Student (Universal Score Variance)	0.9093	0.8948	0.8495	0.8845	48.38%
Construct	0.2049	0.3148	0.0570	0.1928	***10.54%***
Rater	0.00048	0.0002	0.001	0.0005	0.27%
Construct*Rater	0.014	0.014	0.365	0.13	***7.1%***
Student*Construct	0.15	0.1475	0.163	0.1531	8.37%
Student*Rater	0.25	0.277	0.318	0.28	15.31%
Student*Rater*Construct	0.181	0.189	0.192	0.1873	10.24%
Total Variance	1.70968	1.8373	1.9455	1.8282	100%
G Coefficient	***0.788***	***0.775***	***0.747***		

Table 7: Estimates of Variance Components for G Study Two Facets Fully Crossed Design

(Grade 5)

Source of Variation	Estimated Variance Components				
	Sample A (N= 227)	Sample B (N=231)	Sample C (N=224)	Overall Average	
Student (Universal Score Variance)	0.7175	0.5450	0.5760	0.61	36.9%
Construct	0.1130	0.0520	0.1923	0.12	***7.26%***
Rater	0.002	0.004	0.001	0.003	0.18%
Construct*Rater	0.12	0.46	0.45	0.34	***20.6%***
Student*Construct	0.18	0.20	0.15	0.18	10.89%
Student*Rater	0.18	0.22	0.23	0.21	12.7%
Student*Rater*Construct	0.17	0.204	0.185	0.19	11.49%
Total Variance	1.4825	1.685	1.7843	1.653	100%
G Coefficient	***0.763***	***0.677***	***0.709***		

Table 8: Estimates of Variance Components for G Study Two Facets Fully Crossed Design

(Grade 8)

Source of Variation	Estimated Variance Components				
	Sample A (N=235)	Sample B (N=235)	Sample C (N=242)	Overall Average	
Student (Universal Score Variance)	1.0039	.9854	0.9388	0.976	32.3%
Construct	0.0668	0.0956	4.2341	1.465	***48.49%***
Rater	0.001	0.002	0.000	0.001	0%
Construct*Rater	0.026	0.019	0.154	0.066	***2.18%***
Student*Construct	0.092	0.132	0.138	0.12	3.92%
Student*Rater	0.270	0.257	0.204	0.244	8.1%
Student*Rater*Construct	0.150	0.153	0.142	0.148	4.9%
Total Variance	1.6079	1.644	5.8109	3.021	100%
G Coefficient	***0.821***	***0.809***	***0.820***		

Discussions

The data used for this analysis contains only one writing task for each student. Not only it is inappropriate, but also impossible to assess the reliability or dependability of this assessment within traditional reliability analytical framework. However, the assessments were designed focusing on the constructs underlying domains tested. Meanwhile, scoring guides were developed centering the constructs comprised of dimensions for raters to consider. Raters were trained to follow the construct-centered-scoring guide. To assess the dependability of the assessments involved, by incorporating contextual information provided in the scoring guide, we were able to decompose the one complex writing task into a number of smaller constructs, which were comprised more homogeneous dimensions. The rationale guiding the approach centered focus on the constructs that were intended to be measured by the domain tested. Thus, we differentiate this approach from traditional reliability analytical approach.

Looking at the results of generalizability analysis obtained using grade 3, 5 and 8 data, we found that high generalizability or reliability is achievable using construct centered reliability analytical approach where there is a high degree of fit between the substantive expectations generated from the test specialist's understanding of the construct-as realized by the measurement procedure-and observation (Nichols and Smith, 1998).

Educational or Scientific Importance of Study:

Instead of increasing number of writing prompts or writing tasks per student, we incorporated contextual information when we performed generalizability analysis by centering our focus on underlying constructs of domain tested, thus defining more variance systematically. The evidence from this study support the need for reliability studies to incorporate theories of learning and performance and that this incorporation has practical implications for the use of different assessment practices. Test developers have an obligation to accommodate different sets of substantive assumptions into the evaluation of measurement procedure's reliability (Nichols & Smith, 1998). The construct centered reliability analytical model serves an alternative to the traditional reliability analytical model when assessing the dependability of cognitively complex

task. This study provides an empirical example of a framework for contexualizing the interpretation of reliability data, which is critical to validate performance assessment.

References

Cronbach, L. J. (1988). Five perspectives on validation argument. In H. Wainer & H. Braun (Eds.), Test validity (pp. 3-17). Hillside, NJ: Erlbaum.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). The dependability of behavioral measurements: Theory of generalizability for scores and profiles. New York: Wiley.

Jiang, Y.H., Smith, P.L., & Nichols, P.D. (1997) Error Sources influencing reliability or generalizability of performance assessment: A meta-analysis. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Koretz, D., McCaffrey, D., Klein, S., Bell, R., & Stecher, B. (1992). The reliability of scores from the 1992 Vermont portfolio assessment program (Report No. TM 019 641). (ERIC Document Reproduction Service No. ED 355 284).

Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational measurement (3rd ed., pp. 13-103). Washington, DC: American Council on Education & National Council on Measurement in Education.

Nichols, P. D., & Smith, P. L. (1998). Contextualizing the Interpretation of Reliability Data. Educational Measurement: Issues and Practice 24-36, Fall, 1998

Nichols, P. D., & Surgrue, B. (1997, August). NAEP redesign: Short forms and construct-centered test development. Paper presented at the National Center for Education Statistics. Washington, DC.

Romberg, T. A., & Wilson, L. D. (1995). In T.A. Romberg (Ed.), Mathematics assessment and evaluation: Imperatives for mathematics educators (pp. 12-28). Albany, NY: State University of New York Press.

Shavelson, R. J., Gao, X., & Baxter, G. P. (1993). Sampling variability of performance assessments. Journal of Educational Measurements, 30(3), 215-232.

Wiggins, G. P. (1993). Assessing student performance: Exploring the purpose and limits of testing: San Francisco: Jossey-Bass.

Appendix

TABLE 2: Demographic Characteristics of Students Tested in English for Grade 3, 5 and 8

Grade	Gender		Ethnicity							English Language Fluency			Free/Reduced Lunch			Total N
	Male	Female	Asian	African American	Filipino	Hispanic	American Indián	Pacific Islander	White	English Only	Fluent English Proficiency	English Language Learner	Full Price	Reduced Price	Free	
3	51.2%	48.8%	15.7%	26.7%	4.0%	26.5%	0.3%	3.0%	23.9%	64.2%	7.7%	28.1%	34.4%	58.0%	7.6%	5488
5	51.1%	48.9%	14.8%	23.4%	3.0%	35.7%	0.5%	2.3%	20.4%	54.5%	17.6%	27.9%	31.4%	60.0%	8.2%	6009
8	50.7%	49.3%	17.2%	20.9%	3.7%	36.2%	0.4%	2.1%	19.6%	50.0%	25.0%	25.0%	32.7%	59.9%	7.4%	5833

TABLE 4: Constructs Intended To Be Measured and Dimensions Within the Constructs By Grade 3 Writing Prompts								
Constructs	**Dimension Index**	**Dimension Descriptions**	**6**	**5**	**4**	**3**	**2**	**1**
Incidence	IQ1	Coherent and engaging story	x					
	IQ2	Moves toward central moment with drama	x					
	IQ3	Tells readers what they need to know to understand what happened	x					
	IQ4	Readers can infer the incident's significance to the writer	x					
	IQ5	Sensory descriptions	x					
	IQ6	Narrating specific action	x					
	IQ7	Creating dialogues	x					
	IQ8	Slowing the pace to elaborate central moment in the incident	x					
	IQ9	Creating suspense or tension	x					
	IQ10	Including the element of surprise	x					
	IQ11	comparing or contrasting other scenes or people	x					
	IQ12	Less drama than a "6"		x				
	IQ13	Structurally more predicatble than a "6"		x				
	IQ14	Less focused, especially toward the end		x				
	IQ15	Uses a narrower range of narrative strategies		x				
	IQ16	Lacks the authority of a "5"			x			
	IQ17	May be momentary digreessions			x			
	IQ18	Story may be smoothly told yet unrealized dramatically.			x			
	IQ19	Limited use of narrative strategies			x			
	IQ20	Related specific incident				x		
	IQ21	Story competently told				x		
	IQ22	Brief				x		
	IQ23	Flat, unfocused				x		
	IQ24	May be series of loosely connected events				x		
	IQ25	Very limited use of narrative strategies				x		
	IQ26	May fail to focus on an incident					x	
	IQ27	Incohesion. May tell an incident without orienting context or significance					x	
	IQ28	Usually quite brief.					x	
	IQ29	If longer, may be rambling, fragmentary, or without details.					x	
	IQ30	Contains omissions, erratic jumps in time or place, or breakdowns					x	
	IQ31	May refer to an incident without identifying it specifically						x
	IQ32	May only imply the incident						x
	IQ33	May point to an incident without developing it conclusively						x
	IQ34	Writer may focus on others instead of him/self						x
Context	CQ1	Writer locates incident in a particular setting and orients reader to scene, people and events	x					
	CQ2	Carefully chosen details used to develop the scene or the people	x					
	CQ3	Considerable space devoted to orienting readers, describing the scene and people, and providing back-ground or context for the incident but not at the expense of a well-told incident	x					
	CQ4	Balance achieved between static context and dramatic narrated incident	x					
	CQ5	Appropriate and adequate context as in a 6-point essay		x				
	CQ6	Context does not dominate at the expense of the incident		x				
	CQ7	Adequate to orient readers to the incident			x			
	CQ8	May devote too much space to context while neglecting the narrative				x		
	CQ9	May begin abruptly without necessary orientation				x		
	CQ10	Context is limited or even missing (1)					x	
	CQ11	Context is limited or even missing (2)						x

Constructs	Dimension Index	Dimension Descriptions	6	5	4	3	2	1
Voice and style	VQ1	Authentic voice	x					
	VQ2	Reveals writer's attitude toward the incident	x					
	VQ3	Well chosen details	x					
	VQ4	Appropriate words	x					
	VQ5	Graceful, varied sentences	x					
	VQ6	Often includes word play and imagery	x					
	VQ7	Engages the reader from the beginning and moves to a satisfying closure	x					
	VQ8	Competent stylistically		x				
	VQ9	May lack the grace, surprise, or sparkle of a "6"		x				
	VQ10	Begins engagingly and closes in a satisfying way		x				
	VQ11	The voice of an earnest story-teller			x			
	VQ12	Predicatable sentences and word voice			x	x		
	VQ13	Writer relates incident in an uninvolved way				x		
	VQ14	Writer does not seem to be seeing the incident as it happened				x		
	VQ15	Minimal evidence of personal involvement					x	
	VQ16	Writer doe not seem to be relating specific details about the incident					x	
	VQ17	Sentences may be too short or long in a disorderly way					x	
	VQ18	Writer communicates little or no evidence of personal involvement in the incidence						x
Significance	SQ1	Reveals by statement of implication, why the incident was important to the writer	x					
	SQ2	Significance may be apparent in the writer's insights at the time of the incident or in reflections from his/her present perspective.	x					
	SQ3	Insights/reflections may appear integrated into the narration or in the conclusion	x					
	SQ4	Reflections may be humorous	x					
	SQ5	Signifance either implied or stated clearly, through remembered or present reflections		x				
	SQ6	Reflections not as perceptive as 6-point essay, but not superficial		x				
	SQ7	Less well integrated as 6-point essay; often at end of essay		x				
	SQ8	Either implied or stated			x			
	SQ9	Reflection not as insightful as 5-point essay			x			
	SQ10	Reflection may seem tacked on at the end of the essay			x			
	SQ11	Implied or briefly stated				x		
	SQ12	Gives readers an idea why the incident was memorable				x		
	SQ13	Reflections not especially insigniful				x		
	SQ14	Few, if any, reflections					x	
	SQ15	Reflections may seem superficial					x	
	SQ16	Little or no significance implied or stated						x

TABLE 5: Constructs Intended To Be Measured and Dimensions Within the Constructs By Grade 8 Writing Prompts								
Constructs	**Dimension Index**	**Dimension Descriptions**	**6**	**5**	**4**	**3**	**2**	**1**
Presentation of the situation	PQ1	Clearly defines, identifies or describes the situation to be speculated about	x					
	PQ2	Situation is presented fully and precisely, but does not dominate the essay at the expense of speculation	x					
	PQ3	Writer limits the occasion appropriately; readers' attention is focused on just those aspects the writer will speculate about	x					
	PQ4	The presentation of the situation grounds and focuses the entire essay	x					
	PQ5	Reader is immediately engaged	x					
	PQ6	Language is conrete, rich in sensory detail	x					
	PQ7	Writer uses narrative and descriptives strategies	x					
	PQ8	Writer acknowledges readers; concerns/questions.	x					
	PQ9	Writer convinces readers of the plausibility of the speculation.	x					
	PQ10	For real-world situations, the writer acknowledges reader's experience or familiarity with a situation, then builds on this to focus reader attention on a comparable situation	x					
	PQ11	Writer establishes authority by consistently demonstrating broad knowledge and clear understanding of the situation	x					
	PQ12	Situation is clearly defined, but with less elaboration than a "6"		x				
	PQ13	Situation does not dominate the essay at the expense of speculation		x				
	PQ14	Writer limits and focuses the occasion, but with less panache than a "6"		x				
	PQ15	Writer's knowledge and understanding of the situation is clear throughout, and a sense of confidence and authority is maintained		x				
	PQ16	Language lacks only the vividness and impact of a "6"		x				
	PQ17	Relies on a narrower range of strategies for presenting the situation than a "6"		x				
	PQ18	Situation presented with less assuredness than a "5" or "6"			x			
	PQ19	Situation may tend to dominate the essay			x			
	PQ20	Not as clearly focused, or may lack detail or specifically or a "5" or "6"			x			
	PQ21	The presentation of the situation is adequate to orient readers to the proposed causes or effects			x			
	PQ22	Essay contains some explicit speculation			x			
	PQ23	Writing presents a situation				x		
	PQ24	Situation may either be brief or may dominate the essay				x		
	PQ25	Writer may paraphrase the prompt rather than define the situation				x		
	PQ26	Writer may not clearly establish the boundaries of the situation				x		
	PQ27	Writer may not seem to fully understand the situation				x		
	PQ28	Commonplace language				x		
	PQ29	Limited use of strategies				x		
	PQ30	Writer may not acknowledge readers				x		
	PQ31	Writer may attempt to construct a situation but, because of omissions, erratic jumps in time or place or breakdowns in cohesion, will not establish focus.					x	
	PQ32	Situation may dominate the essay					x	
	PQ33	Essay may include no occasion, begining abruptly with a list of causes or effects					x	
	PQ34	Writer exhibits only minimal understanding of the situation					x	
	PQ35	If there is a situation, it will be very brief and devoid of speficity or conretness						x
	PQ36	Essay may point vaguely to a situation without focusing or establishing boundaries						x

Constructs	Dimension Index	Dimension Descriptions	6	5	4	3	2	1
Logic and relevance of causes	LQ1	Proposes causes and effects are clearly related to the situation	x					
	LQ2	Writer uses imaginative, investigative, inventive arguments to convince readers of the logic of the speculations	x					
	LQ3	Multiple perspectives/possibilities are considered	x					
	LQ4	Writer stretches imagination to take ideas as far as possible, are convinced that the writer's speculations are plausible and appropriate to the situalations as defined	x					
	LQ5	Writer maintains focus by establishing and continually developing the close relationship between the particular situation and the causes/effects that might arise from it	x					
	LQ6	Writer is continually aware of reader's needs	x					
	LQ7	Writer may employ some of the following strategies:	x					
	LQ8	Building a succession of causes of effects, each changing the complex:	x					
	LQ9	Writer establishes, maintains, and developing a plausible relationship between the situation and each of the processed causes or effects	x					
	LQ10	The speculation is naturally linked to the situation.	x					
	LQ11	Transitions skillfully keep the reader grounded both in the relationship between the situation and the proposed causes and effects and in the logical development and progression of the speculation itself	x					
	LQ12	Writer uses the transitions to carry the reader along with the methodlogical development of the argument	x					
	LQ13	Writer weaves together facts, opinions & projections to create and develop convincing reasons	x					
	LQ14	Proposed causes or effects are linked naturally to the defined situation		x				
	LQ15	Writer concjectures persuasively for possible causes or effects		x				
	LQ16	Speculations are serious and logical, lacking only the freshness and imagination of a "6".		x				
	LQ17	Obviously statements about probable causes and effects to speculations that are not entirely predictable.		x				
	LQ18	Essay may lack the clarity of focus, the continuity, or the growing insight and fullness of a "6".		x				
	LQ19	Writer has a consistent awareness of audience		x				
	LQ20	Speculations are insightful, but not as probing as a "6"		x				
	LQ21	Writer keeps the reader grounded in both situation and the speculation and the speculation, although not as consistently as a "6"		x				
	LQ22	Writer establishes a connection between the situation and the postulated causes or effects, but may not maintain this connection as explicitily or effectively as "5" or "6".			x			
	LQ23	Speculations characterized by thoughtfulness rather than inventiveness			x			
	LQ24	Proposed causes or effects may be logical but predicatable			x			
	LQ25	Speculations may be connected			x			
	LQ26	Acknowledgment of readers not as evident as in a "5" or "6"			x			
	LQ27	Speculations are at least tangentially relevant.				x		
	LQ26	Writer may tend to list a series of causes or effects rather than develop them or ground them in the situation.				x		
	LQ29	Little effort to convince the reader by developing a logical cause-effect relationship.				x		
	LQ30	Speculations may seem obvious, superfiicial, or predicatable.				x		
	LQ31	Proposed causes/effects arise from or are appropriate to the situation, but may seem tangential and not grounded as firmly in the situations.				x		
	LQ32	Little conscious awareness of the reader.				x		
	LQ33	Essay may have a meandering quality				x		
	LQ34	May be only one minimally developed cause/effect				x		
	LQ35	If there is a situation, the speculations may be either brief or meandering/unfocused					x	
	LQ36	Little evidence of any logical organization					x	
	LQ37	Some of the proposed causes/effects may seem illogical or unrelated to the situation					x	
	LQ38	Little connection between the situation and the speculations					x	
	LQ39	If there are speculations, they are brief and superficial attempts at prediction rather than considered exploration of possibilities						x
	LQ40	No evidence of a logical connection between the situation and the causes or effects.						x

Constructs	Dimension Index	Dimension Descriptions	6	5	4	3	2	1
Elaboration of argument	EQ1	Essay provides substantial elaboration.	x					
	EQ2	Reader is convinced that the writer's conjectures are valid for the situations	x					
	EQ3	Writer uses carefully chosen evidence that is logically and fully developed.	x					
	EQ4	Reader is convinced of both the logic adn the authenticity of the proposed cause/effect.	x					
	EQ5	Strategies used to develop arguments:	x					
	EQ6	Writer may mention several possible causes/effects, developing and linking them.	x					
	EQ7	Writer may only mention one cause/effect, building it fully adn examining it closely from a variatey perspectives.	x					
	EQ8	Writer makes a full and convincing argument for at least one postulated cause and one postulated effect.	x					
	EQ9	Writer engages in extended, thoughtful speculation		x				
	EQ10	Writer uses effective arguments to convince the reader of the logic adn validity of the speculations.		x				
	EQ11	Writer chooses evidence that is relevant and convincing.		x				
	EQ12	Supporting evidence is more predictable than a "6".		x				
	EQ13	Supporting details are relevant and convincing; not as richly developed as a "6".		x				
	EQ14	Writer offers less persuasive evidence for the validity of the proposed causes/effects.			x			
	EQ15	Essay exhibits some internal logic and an over-all sense of organization.			x			
	EQ16	Essay may not show a consistent relationship between the situation and cause/effects.			x			
	EQ17	May be some irrelevant details.			x			
	EQ18	Elaboration is limited - perhaps to a brief explanation of one cause/effect or a listing of several with minimal development.				x		
	EQ19	Essay lacks consistency of development of details.				x		
	EQ20	Sequence and organizational pattern seem unclear.				x		
	EQ21	Essay may seem generally competent and the speculations interesting				x		
	EQ22	Little elaboration, often merely listing.					x	
	EQ23	Essay often brief.					x	
	EQ24	Little development either of the situation or of cause/effects.					x	
	EQ25	May be extended generalized rambling.					x	
	EQ26	May merely list causes/effects without support of argument.					x	
	EQ27	Some details may be irrelavant and unconnected to either the speculations or the situation; Essay contains little or no argument or effort to persuade the reader.					x	
	EQ28	Little or no elaboration of either the situation or of the causes/effects.						x
	EQ29	Speculations, if presented at all, are not argued.						x
	EQ30	Rarely is there any sense of the reader.						x
	EQ31	Essay is brief; Essay is often not coherent.						x

TABLE 6: Constructs Intended To Be Measured and Dimensions Within the Constructs By Grade 5 Writing Prompts								
Constructs	*Dimension Index*	*Dimension Descriptions*	*6*	*5*	*4*	*3*	*2*	*1*
Identification of subject	IQ1	Clearly defined or identified	x	x				
	IQ2	Clearly identified but rarely defined			x			
	IQ3	The subject is identified, but in a very general rather than specific way				x		
	IQ4	More than one subject may be introduced					x	
	IQ5	May or may not indicate a subject or subjects						x
Context	CQ1	Clearly located in a specific physical or psychological context	x					
	CQ2	Located in a specific physical context		x				
	CQ3	Located in a context, but the context may either be very scant, or may dominate the essay at the expense of the subject			x			
	CQ4	The subject is located in a context, but may either be very minimal, or it may dominate the essay at the the expense of the subject				x		
	CQ5	The context is minimal; when there are multiple subjects, context may be completely lacking					x	
	CQ6	No context is provided						x
Observational stance	OQ1	The writer's role as observer, rather than active participants, is clearly established. This relationship between the writer as eyewitness and the subject is developed, and maintained throughout the essay.	x					
	OQ2	The writer conveys a sense of discovery that has resulted from his/her observations	x					
	OQ3	The relationship between the writer as eyewitness and the subject is well established		x				
	OQ4	The relationship between the writer and the subject is less clear than in five and six-point essays			x			
	OQ5	The relationship between the writer and the subject is inconsistent or may tend to be autobiographical rather than observational				x		
	OQ6	The observer is central, often evaluative, and the writing is autobiographical					x	
	OQ7	The observer is central, the subject, when discernible, may be mentioned only once or twice						x
Presentation of the experience	PQ1	Writing is rich in sensory detail and concrete images	x					
	PQ2	The reader can recognize the meaning, impact, and importance that the experience had for the writer	x					
	PQ3	The writer skillfully uses a variety of presentational strategies such as dialogue or monologue, narration, specific action/behavior, shifting perspectives, and/or metaphoric language to recreate his/her perception for the reader	x					
	PQ4	Throughout the essay, the writer exhibits a controlled awareness of the scence itself, of his/her own purpose, and of the developing effect of the essay on the reader	x					
	PQ5	Provides sufficient detail for the reader to see the scene clearly		x				
	PQ6	Recognize the meaning and importance that the experience had for the writer		x				
	PQ7	The writer uses a number of presentational stragtegies, but fewer or less effectively than a six-point essay.		x				
	PQ8	Throughout the essay, the writer exhibits a fairly well controlled awareness of the scence itself, of his/her own purposes, and of the developing effect of the essay on the reader		x				
	PQ9	Provides some detail, but it is often skimpy and/or unelaborated			x			
	PQ10	May not convey the meaning or importance for the writer			x			
	PQ11	The writer uses several pretational strategies, but with little elaboration			x			
	PQ12	The writer shows little awareness of the reader			x	x		
	PQ13	Writing is characterized by generalities rather than specific details				x	x	
	PQ14	Does not convey the meaning or importance of the experience				x	x	x
	PQ15	The writer uses very few presentational strategies				x		
	PQ16	The writer rarely shows awareness of the reader					x	
	PQ17	Information about the subject is garbled or lacking						x
	PQ18	The writer shows no awareness of the reader						x

Summary of Rhetorical Effectiveness Scoring Guide for
AUTOBIOGRAPHICAL INCIDENT

	6	5	4	3	2	1
INCIDENT	• Coherent and engaging story. • Moves toward central moment with drama. • Tells readers what they need to know to understand what happened. • Readers can infer the incident's significance to the writer. • Uses some of the following narrative strategies: • sensory descriptions • narrating specific action • creating dialogues, interior monologues • slowing the pace to elaborate central moment in the incident • creating suspense or tension • including the element of surprise • comparing or contrasting other scenes or people	• Coherent and engaging story. • Less drama than a "6." • Structurally more predictable than a "6." • Less focused, especially toward the end. • Uses a narrower range of narrative strategies.	• Lacks the authority of a "5." • May be momentary digressions • Story may be smoothly told yet unrealized dramatically. • Limited use of narrative strategies.	• Relates specific incident. • Story competently told. • Brief. • Flat, unfocused. • May be series of loosely connected events. • Very limited use of narrative strategies.	• May fail to focus on an incident. • In cohesion. May tell an incident without orienting context or significance. • Usually quite brief. • If longer, may be rambling, fragmentary, or without details. • Contains omissions, erratic jumps in time or place, or breakdowns.	• May refer to an incident without identifying it specifically. • May only imply the incident. • May point to an incident without developing it conclusively. • Writer may focus on others instead of him/self.
CONTEXT	• Writer locates incident in a particular setting and orients reader to scene, people and events • Carefully chosen details used to develop the scene or the people. • Considerable space devoted to orienting readers, describing the scene and people, and providing background or context for the incident--but not at the expense of a well-told incident. • Balance achieved between static context and dramatic, narrated incident.	• Appropriate and adequate context as in a 6 -point essay • Context does not dominate at the expense of the incident.	• Adequate to orient readers to the incident.	• May devote too much space to context while neglecting the narrative • May begin abruptly without necessary orientation.	• Context is limited or even missing.	• Context is limited or even missing.

Summary of Rhetorical Effectiveness Scoring Guide for
AUTOBIOGRAPHICAL INCIDENT (Continued)

	6	5	4	3	2	1
VOICE AND STYLE	• Authentic voice. • Reveals writer's attitude toward the incident. • Well chosen details. • Appropriate words. • Graceful, varied sentences. • Often includes word play and imagery. • Engages the reader from the beginning and moves to a satisfying closure.	• Authentic • oice • Competent stylistically • May lack the grace, surprise, or sparkle of a "6." • Begins engagingly and closes in a satisfying way.	• The voice of an earnest story-teller. • Predictable sentences and word voice.	• Writer relates incident in an uninvolved way. • Writer does not seem to be seeing the incident as it happened. • Predictable sentences and word choice.	• Minimal evidence of personal involvement. • Writer does not seem to be relating specific details about the incident. • Sentences may be too short or long in a disorderly way.	• Writer communicates little or no evidence of personal involvement in the incident.
SIGNIFICANCE	• Reveals by statement of implication, why the incident was important to the writer. • Significance may be apparent in the writer's insights at the time of the incident or in reflections from his/her present perspective. • Insights/reflections may appear integrated into the narration or in the conclusion. • Reflections may be humorous.	• Significance either implied or stated clearly, through remembered or present reflections. • Reflections not as perceptive as 6-point essay, but not superficial. • Less well integrated as 6-point essay; often at end of essay.	• Either implied or stated. • Reflection not as insightful as 5-point essay. • Reflection may seem tacked on at the end of the essay.	• Implied or briefly stated. • Gives readers an idea why the incident was memorable. • Reflections not especially insightful.	• Few, if any, reflections. • Reflections may seem superficial.	• Little or no significance implied or stated.

Summary of Rhetorical Effectiveness Scoring Guide for
OBSERVATIONAL WRITING

	6	5	4	3	2	1
• IDENTIFICATION OF SUBJECT	• Clearly defined or identified.	• Clearly defined or identified.	• Clearly identified but rarely defined.	• The subject is identified, but in a very general rather than specific way.	• More than one subject may be introduced.	• May or may not indicate a subject or subjects.
• CONTEXT	• Clearly located in a specific physical or psychological context.	• Located in a specific physical context.	• Located in a context, but the context may either be very scant, or may dominate the essay at the expense of the subject.	• The subject is located in a context, but may either be very minimal, or it may dominate the essay at the expense of the subject.	• Context is minimal; when there are multiple subjects, context may be completely lacking.	• No context is provided.
• OBSERVATIONAL STANCE	• The writer's role as observer, rather than active participant, is clearly established. This relationship between the writer as eyewitness and the subject is developed, and maintained throughout the essay. • The writer conveys a sense of discovery that has resulted from his/her observations.	• The relationship between the writer as eyewitness and the subject is well established.	• The relationship between the writer and the subject is less clear than in five and six-point essays.	• The relationship between the writer and the subject is inconsistent or may tend to be autobiographical rather than observational.	• The observer is central, often evaluative, and the writing is autobiographical.	• The observer is central. The subject, when discernible, may be mentioned only once or twice.
• PRESENTATION OF THE EXPERIENCE	• Writing is rich in sensory detail and concrete images. • The reader can recognize the meaning, impact, and importance that the experience had for the writer. • The writer skillfully uses a variety of presentational strategies such as dialogue or monologue, narration, specific actions/behavior, shifting perspectives, and/or metaphoric language to recreate his/her perception for the reader. • Throughout the essay, the writer exhibits a controlled awareness of the scene itself, of his/her own purposes, and of the developing effect of the essay on the reader.	• Provides sufficient detail for the reader to see the scene clearly. • The reader can recognize the meaning and importance that the experience had for the writer. • The writer uses a number of presentational strategies, but fewer or less effectively than a six-point essay. • Throughout the essay, the writer exhibits a fairly well controlled awareness of the scene itself, of his/her own purposes, and of developing effect of the essay on the reader.	• Provides some detail, but it is often skimpy and/or unelaborated. • May not convey the meaning or importance for the writer. • The writer uses several presentational strategies, but with little elaboration. • The writer shows little awareness of the reader.	• Writing is characterized by generalities rather than specific details. • Does not convey the meaning or importance of the experience. • The writer uses very few presentational strategies. • The writer shows little awareness of the reader.	• Writing is characterized by generalities rather than specific details. • Does not convey the meaning or importance of the experience. • The writer rarely shows awareness of the reader.	• Information about the subject is garbled or lacking. • Does not convey any meaning or importance of the experience. • The writer shows no awareness of the reader.

Summary of Rhetorical Effectiveness Scoring Guide for

SPECULATION ABOUT CAUSES AND EFFECTS/HISTORY ASSESSMENT

	6	5	4	3	2	1
PRESENTING THE SITUATION	• Clearly defines, identifies or describes the situation to be speculated about. • Situation is presented fully and precisely, but does not dominate the essay at the expense of speculation. • Writer limits the occasion appropriately; readers' attention is focused on just those aspects the writer will speculate about. • The presentation of the situation grounds and focuses the entire essay. • Reader is immediately engaged. • Language is concrete, rich in sensory detail. • Writer uses narrative and descriptive strategies. • Writer acknowledges readers; concerns/questions. • Writer convinces readers of the plausibility of the speculation. • For real-world situations, the writer acknowledges reader's experience or familiarity with a situation, then builds on this to focus reader attention on a comparable situation. • Writer establishes authority by consistently demonstrating broad knowledge and clear understanding of the situation.	• Situation is clearly defined, but with less elaboration than a "6". • Situation does not dominate the essay at the expense of speculation. • Writer limits and focuses the occasion, but with less panache than a "6". • Writer's knowledge and understanding of the situation is clear throughout, and a sense of confidence and authority is maintained. • Language lacks only the vividness and impact of a "6". • Relies on a narrower range of strategies for presenting the situation than a "6".	• Situation presented with less assuredness than a "5" or "6". • Situation may tend to dominate the essay. • Not as clearly focused, or may lack detail or specificity of a "5" or "6". • The presentation of the situation is adequate to orient readers to the proposed causes or effects. • Essay contains some explicit speculation.	• Writing presents a situation. • Situation may either be brief or may dominate the essay. • Writer may paraphrase the prompt rather than define the situation. • Writer may not clearly establish the boundaries of the situation. • Writer may not seem to fully understand the situation. • Commonplace language. • Limited use of strategies. • Writer may not acknowledge readers.	• Writer may attempt to construct a situation but, because of omissions, erratic jumps in time or place or breakdowns in cohesion, will not establish focus. • Situation may dominate the essay. • Essay may include no occasion, beginning abruptly with a list of causes or effects. • Writer exhibits only minimal understanding of the situation.	• If there is a situation, it will be very brief and devoid of specificity or concreteness. • Essay may point vaguely to a situation without focusing or establishing boundaries.
LOGIC AND RELEVANCE OF CAUSES	• Proposes causes and effects are clearly related to the situation. • Writer uses imaginative, inventive arguments to convince readers of the logic of the speculations • Multiple perspectives/possibilities are considered. • Writer stretches imagination to take ideas as far as possible. Readers are convinced that the writer's speculations are plausible and appropriate to the situation as defined. • Writer maintains focus by establishing and continually developing the close relationship between the particular situation and the causes/effects that might arise from it. • Writer is continually aware of reader's needs. • Writer may employ some of the following strategies: –"How did" or "If then" patterns of thinking to pose causes and effects and to show the relationship between these conjectures and the situation. • Building a succession of causes or effects, each changing the complex –Showing direct and logical connection between the speculated cause or effect and the situation: how the causes logically create the situation or how the effects arise as a natural consequence of it.	• Proposed causes or effects are linked naturally to the defined situation. • Writer conjectures persuasively for possible, probable, or plausible causes or effects. • Speculations are serious and logical, lacking only the freshness and imagination of a "6". • Writer reaches beyond obvious statements about probable causes and effects to speculations that are not entirely predictable. • Essay may lack the clarity of focus, the continuity, or the growing insight and fullness of a "6". • Writer has a consistent awareness of audience • Speculations are insightful, but not as probing as a "6".	• Writer establishes a connection between the situation and the postulated causes or effects, but may not maintain this connection as explicitly or effectively as a "5" or "6". • Speculations characterized by thoughtfulness rather than inventiveness. • Proposed causes or effects may be logical but predictable. • Speculations may be connected. • Acknowledgment of readers not as evident as in a "5" or "6".	• Speculations are at least tangentially relevant. • Writer may tend to list a series of causes or effects rather than develop them or ground them in the situation. • Little effort to convince the reader by developing a logical cause-effect relationship. • Speculations may seem obvious, superficial, or predictable. • Proposed causes/effects arise from or are appropriate to the situation, but may seem tangential and not grounded as firmly in the situations. • Little conscious awareness of the reader. • Essay may have a meandering quality • May be only one minimally developed cause/effect.	• If there is a situation, the speculations may be either brief or meandering/unfocused. • Little evidence of any local organization. • Some of the proposed causes/effects may seem illogical or unrelated to the situation.	• If there are speculations, they are brief and superficial attempts at prediction rather than considered exploration of possibilities. • No evidence of a logical connection between the situation and the causes or effects.

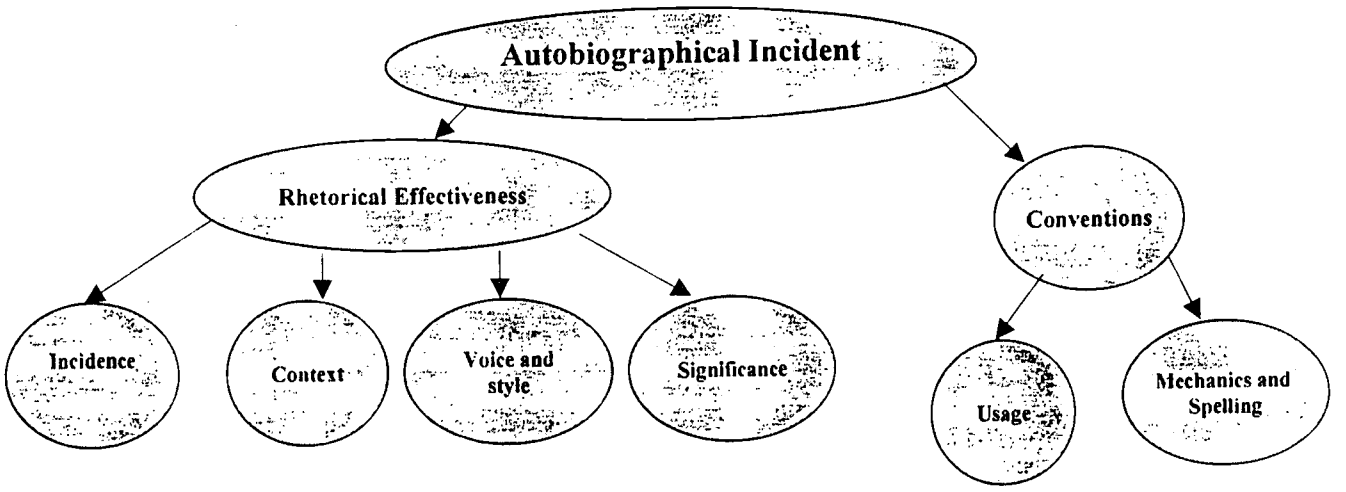
Summary of Rhetorical Effectiveness Scoring Guide for
SPECULATION ABOUT CAUSES AND EFFECTS/HISTORY ASSESSMENT
(Continued)

	6	5	4	3	2	1
LOGIC AND RELEVANCE OF CAUSES (Continued)	–Controlling organizational sequences, such as movement from least to most important cause or effect : from small to large issue; from concrete to abstract; fro m personal to societal causes or effects. • Writer establishes, maintains, and develops a plausible relationship between the situation and each of the proposed causes or effects. • The speculation is naturally linked to the situation. • Transitions skillfully keep the reader grounded both in the relationship between the situation and the proposed causes and effects and in the logical development and progression of the speculation itself. • Writer uses the transitions to carry the reader along with the methodological development of the argument. • Writer weaves together facts, opinions & projections to create and develop convincing reasons.	Writer keeps the reader grounded in both situation and the speculations, although not as consistently as a "6".			• Little connection between the situation and the speculations.	
ELABORATION OF ARGUMENT	• Essay provides substantial elaboration. • Reader is convinced that the writer's conjectures are valid for the situation. • Writer uses carefully chosen evidence that is logically and fully developed. • Reader is convinced of both the logic and the authenticity of the proposed cause/effect. • Strategies used to develop arguments: ☒Citing facts, opinions, projections, and personal experiences or observations (anecdotes) to explain or validate a cause or effect. ☒Elaborating on possibilities arising from proposed causes/effects, showing possible "domino effects" that might determine the direction of a developing situation. ☒Giving specific examples of comparable causes or effects that might have arisen in analogous situations. • Writer may mention several possible causes/effects, developing and linking them. • Writer may only mention one cause/effect, building it fully and examining it closely from a variety of perspectives. • Writer makes a full and convincing argument for at least one postulated cause and one postulated effect.	• Writer engages in extended, thoughtful speculation. • Writer uses effective arguments to convince the reader of the logic and validity of the speculations. • Writer chooses evidence that is relevant and convincing. • Supporting evidence is more predictable than a "6". • Supporting details are relevant and convincing; not as richly developed as a "6".	• Writer offers less persuasive evidence for the validity of the proposed causes/effects. • Essay exhibits some internal logic and an over-all sense of organization. • Essay may not show a consistent relationship between the situation and causes/effects. • May be some irrelevant details.	• Elaboration is limited - perhaps to a brief explanation of one cause/effect or a listing of several with minimal development. • Essay lacks consistency of development of details. • Sequence and organizational pattern seem unclear. • Essay may seem generally competent and the speculations interesting.	• Little elaboration, often merely listing. • Essay often brief. • Little development either of the situation or of causes/effects. • May be extended generalized ramblings. • May merely list causes/effects without support of argument. • Some details may be irrelevant and unconnected to either the speculations or the situation. Essay contains little or no argument or effort to persuade the reader.	• Little or no elaboration of either the situation or of the causes/effects. • Speculations, if presented at all, are not argued. • Rarely is there any sense of the reader. • Essay is brief. Essay is often not coherent.

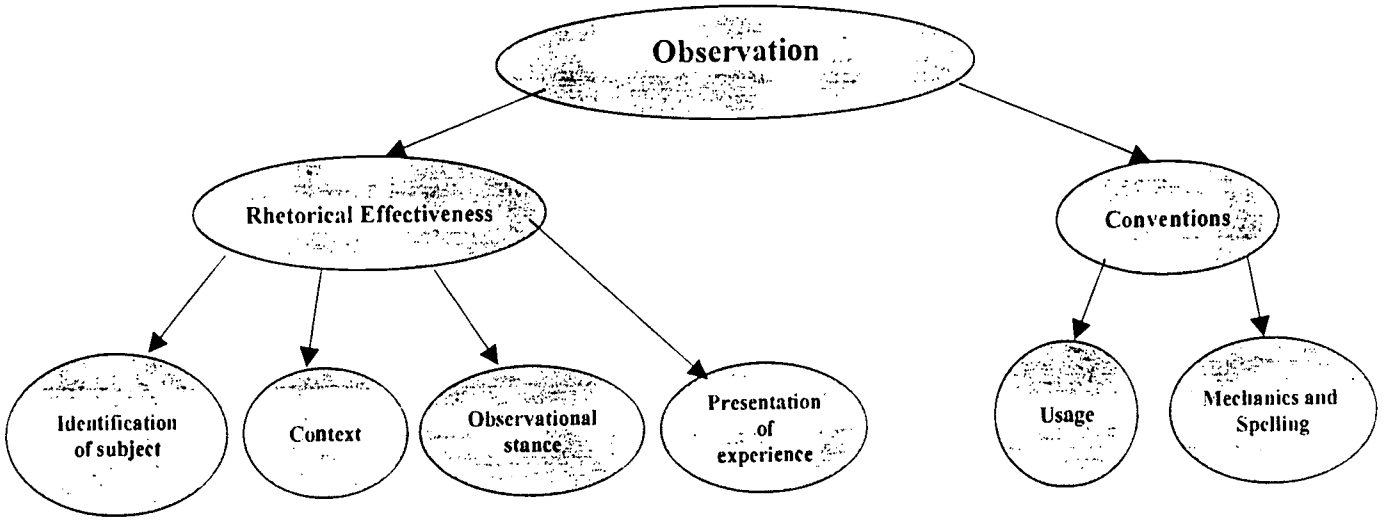
Chart 1

Relationship between Domain Tested and the Underlying Constructs

Grade 3 Domain Tested:



Grade 5 Domain Tested:



Grade 8 Domain Tested:

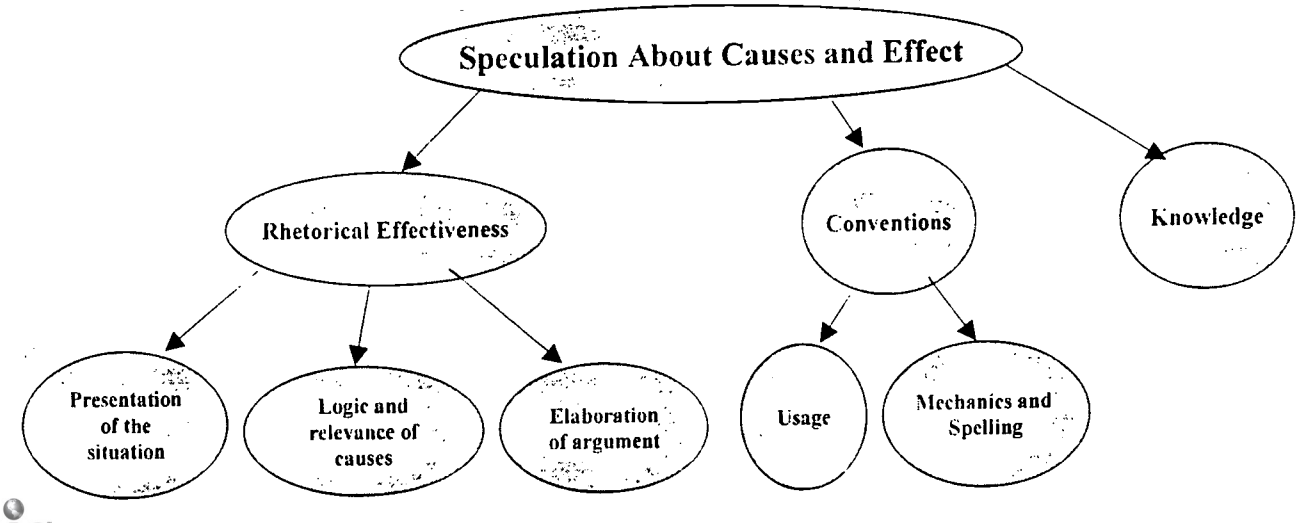


Chart 2

Dimensions indicating the construct of incidence (grade 3)

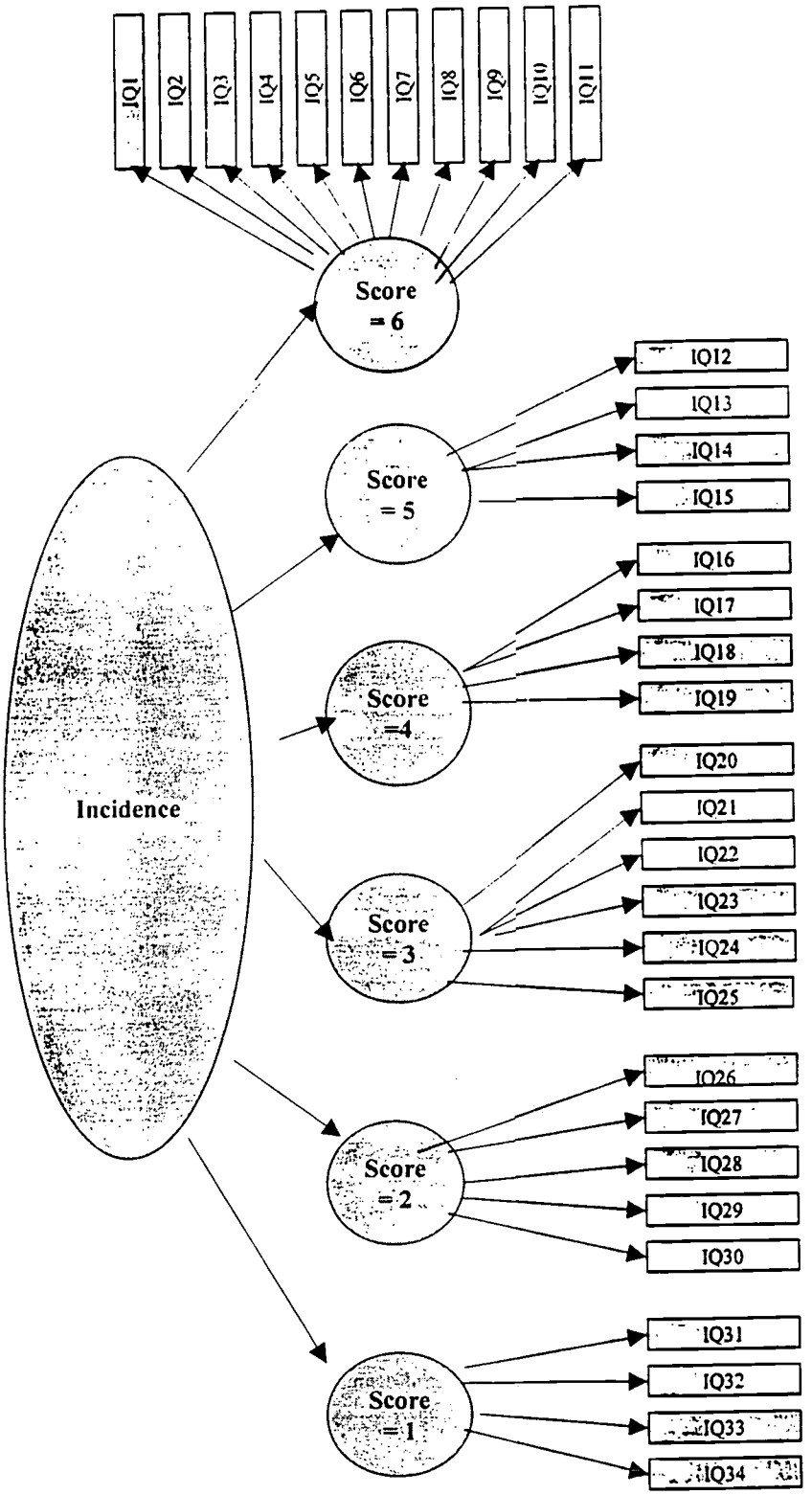


Chart 3

Dimensions indicating the construct of context (grade 3)

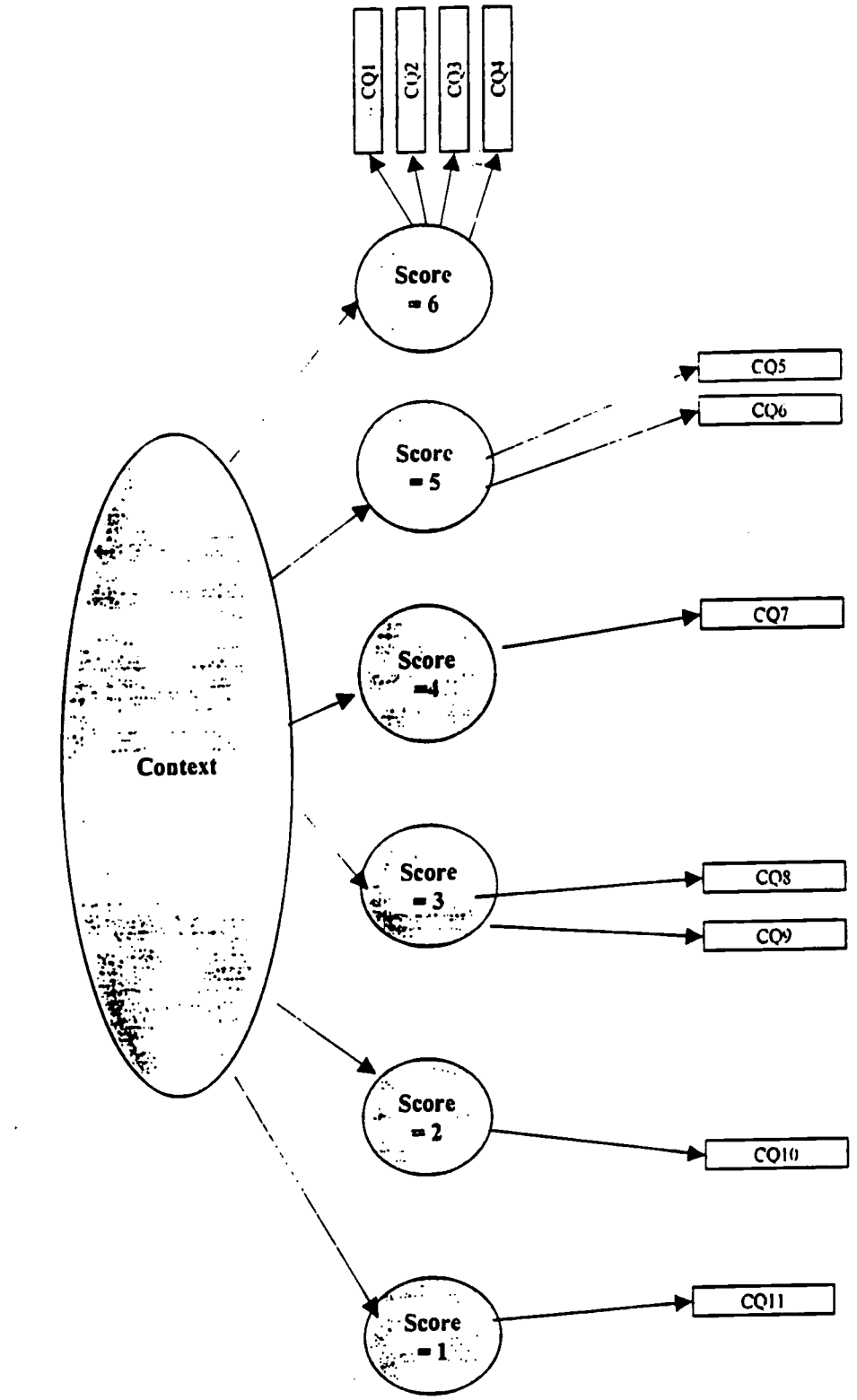


Chart 4

Dimensions indicating the construct of voice and style (grade 3)

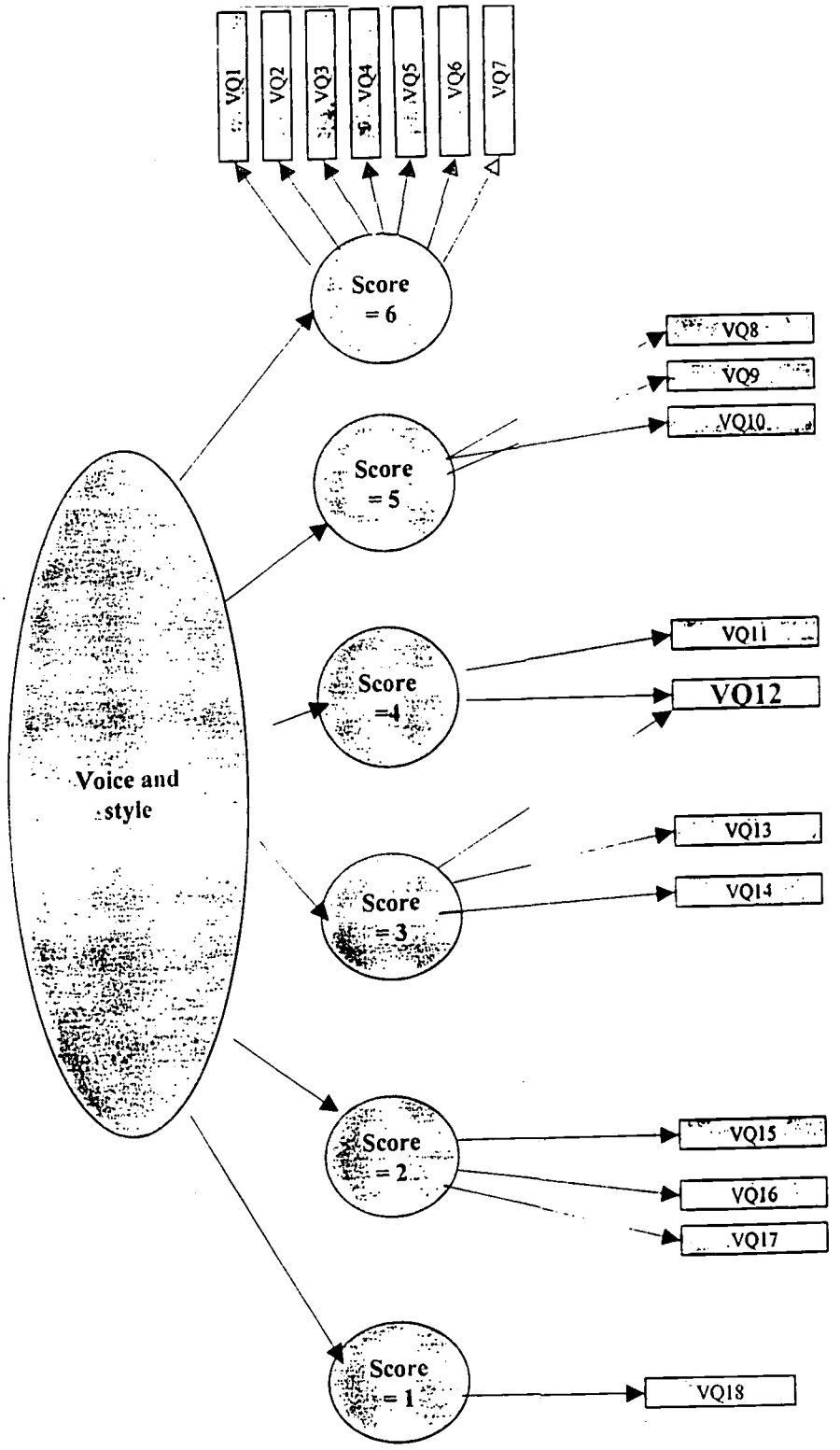


Chart 5

Dimensions indicating the construct of significance (grade 3)

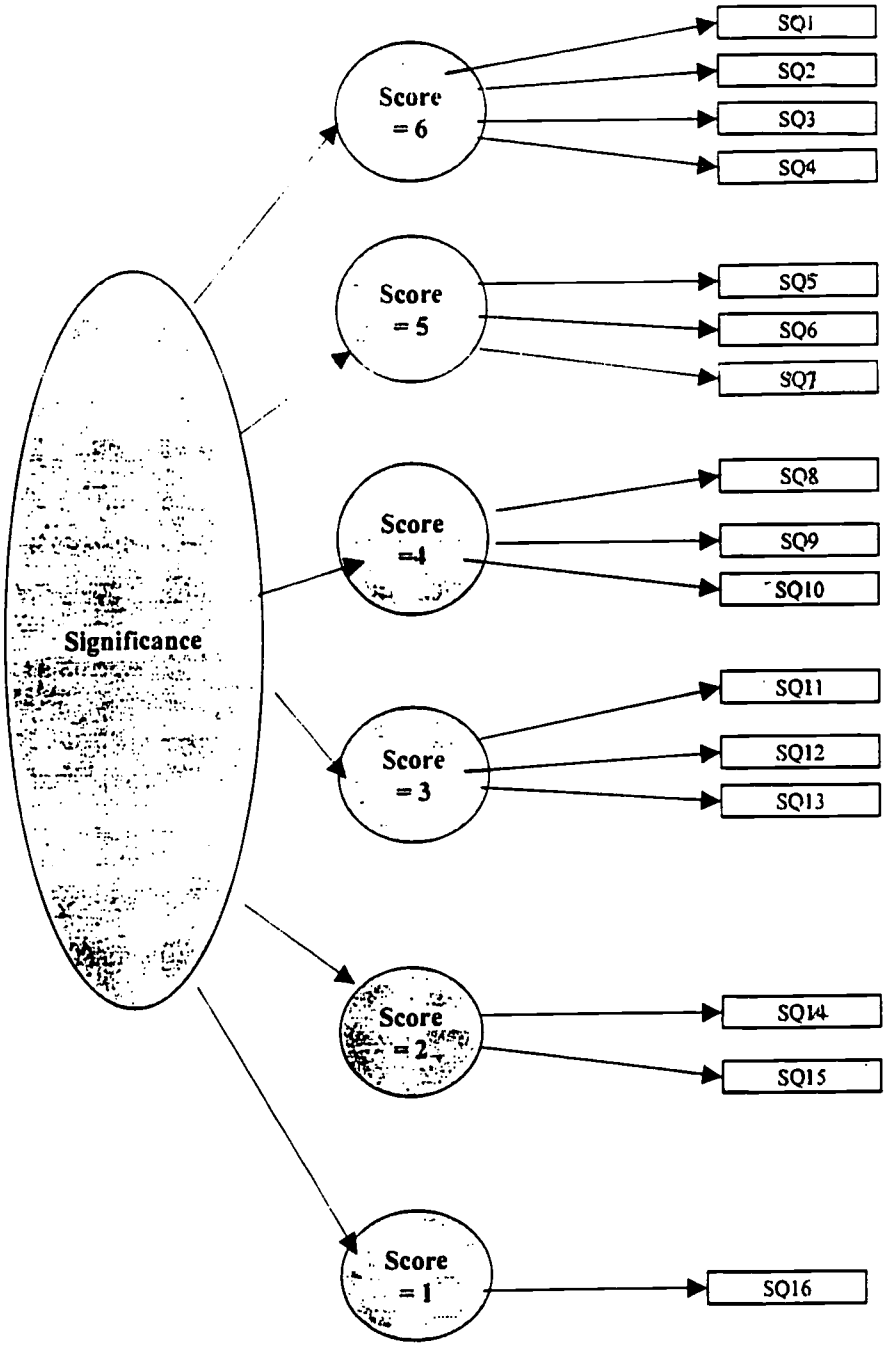


Chart 6

Dimensions indicating the construct of identification of subject (grade 5)

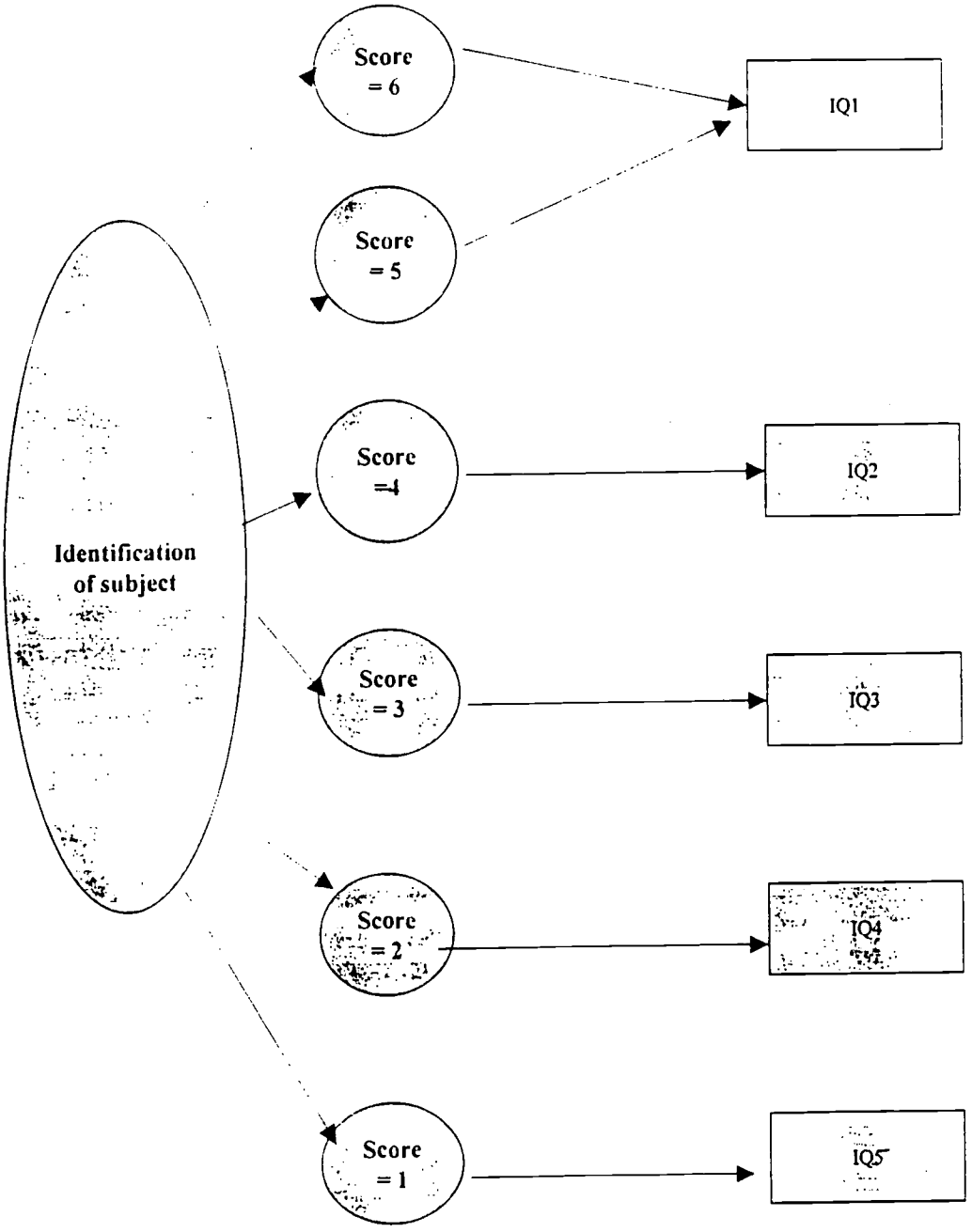


Chart 7

Dimensions indicating the construct of context (grade 5)

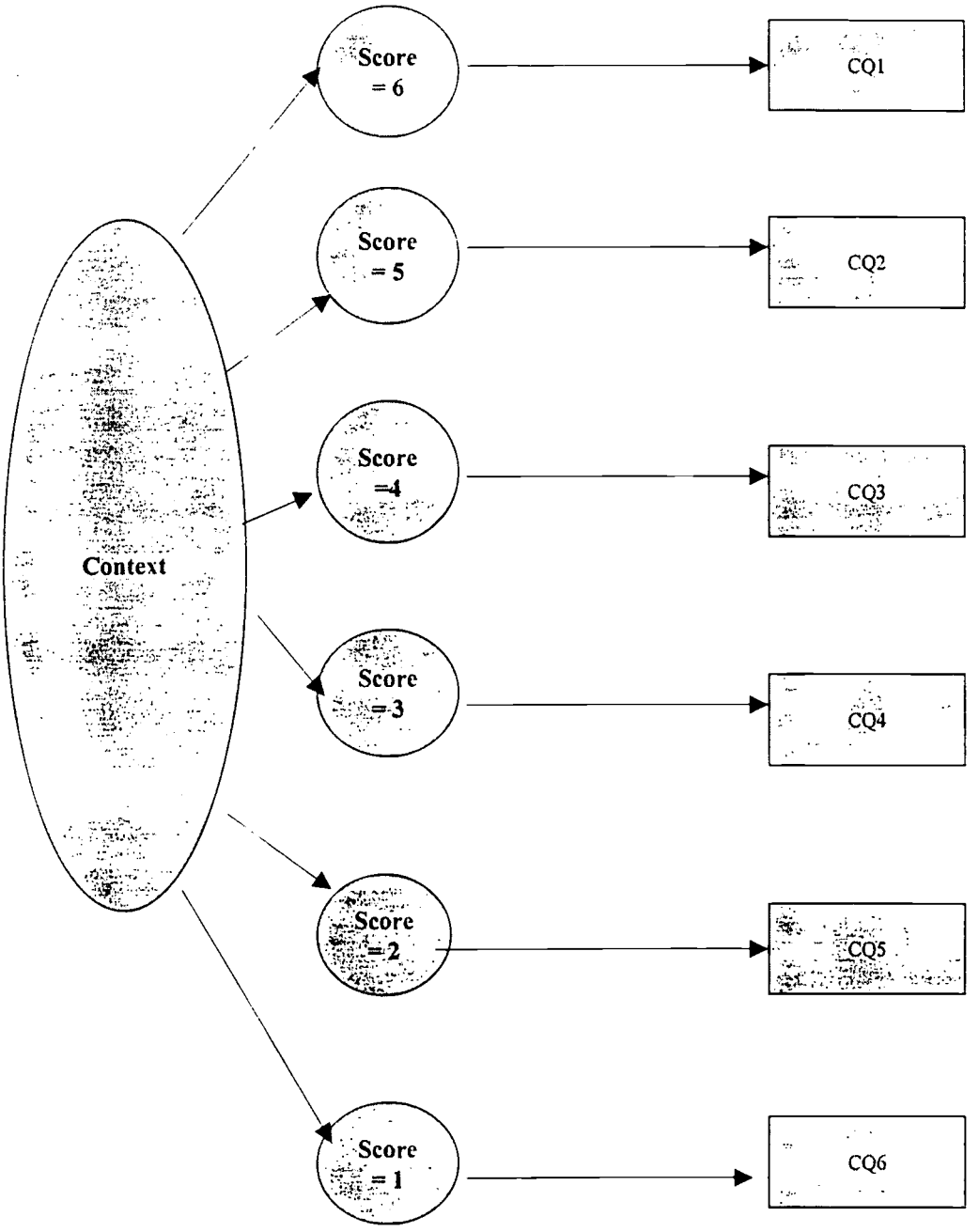


Chart 8

Dimensions indicating the construct of observational stance (grade 5)

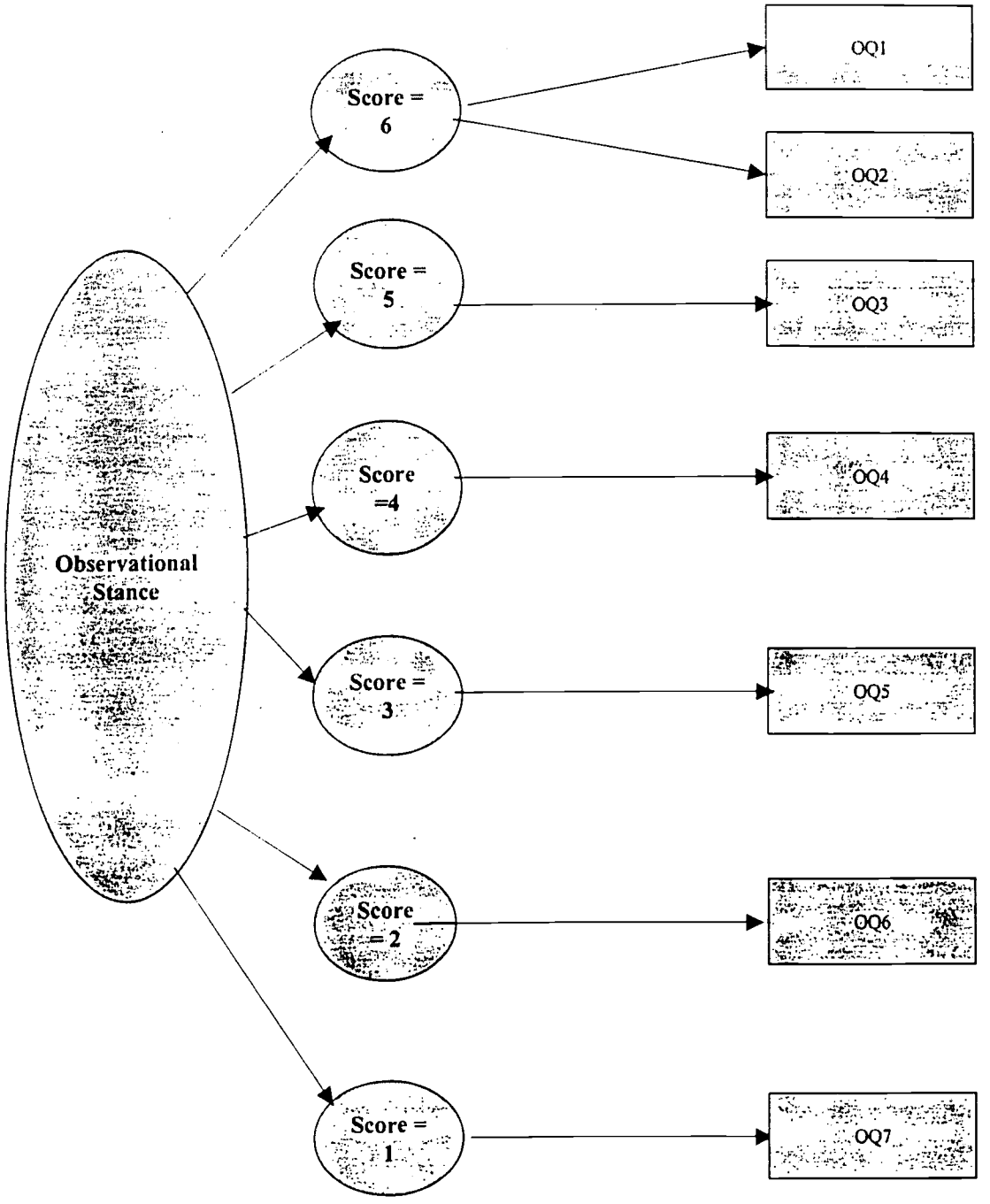


Chart 9

Dimensions indicating the construct of presentation of the experience (grade 5)

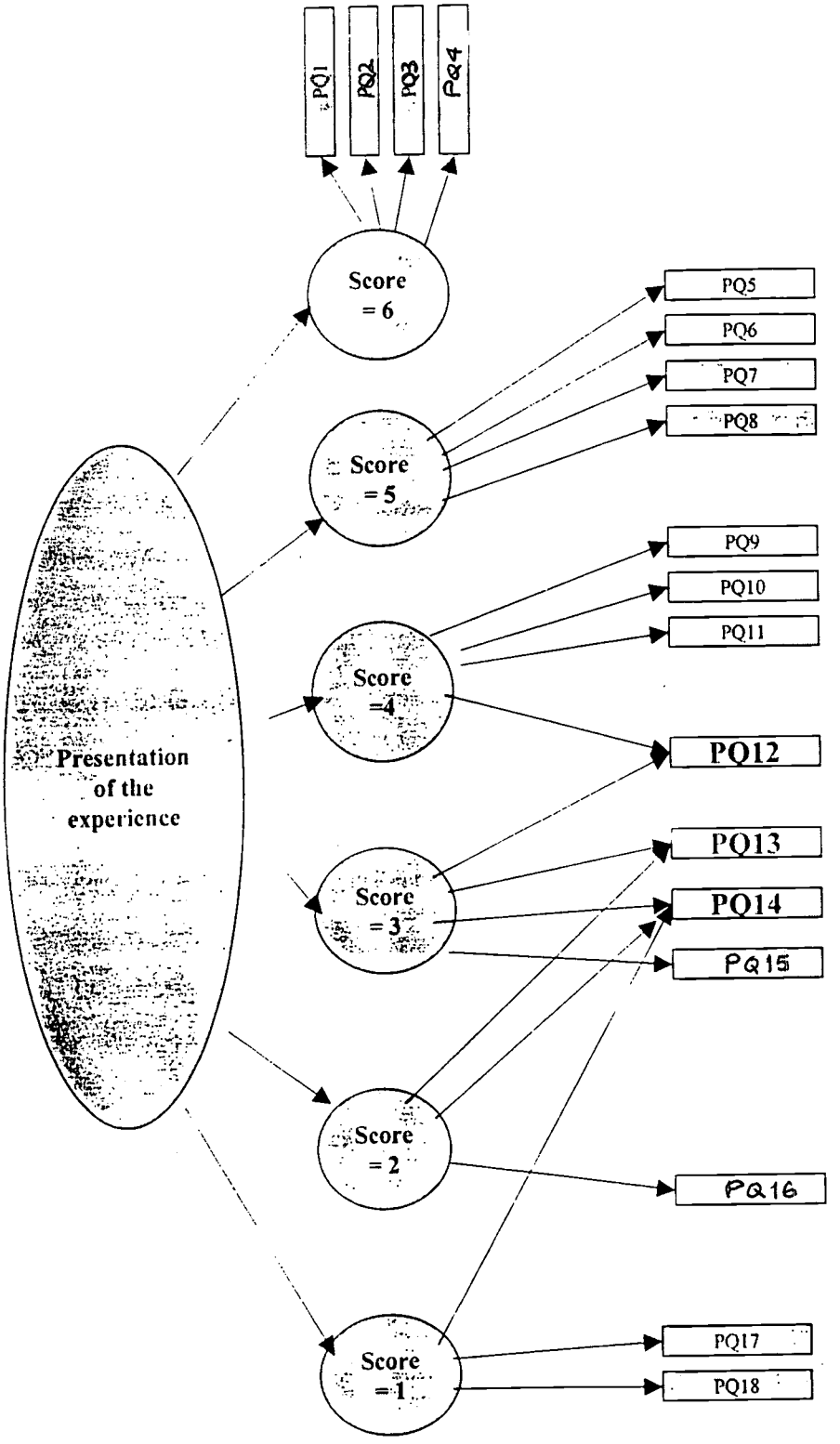


Chart 10

Dimensions indicating the construct of presentation of the situation (grade 8)

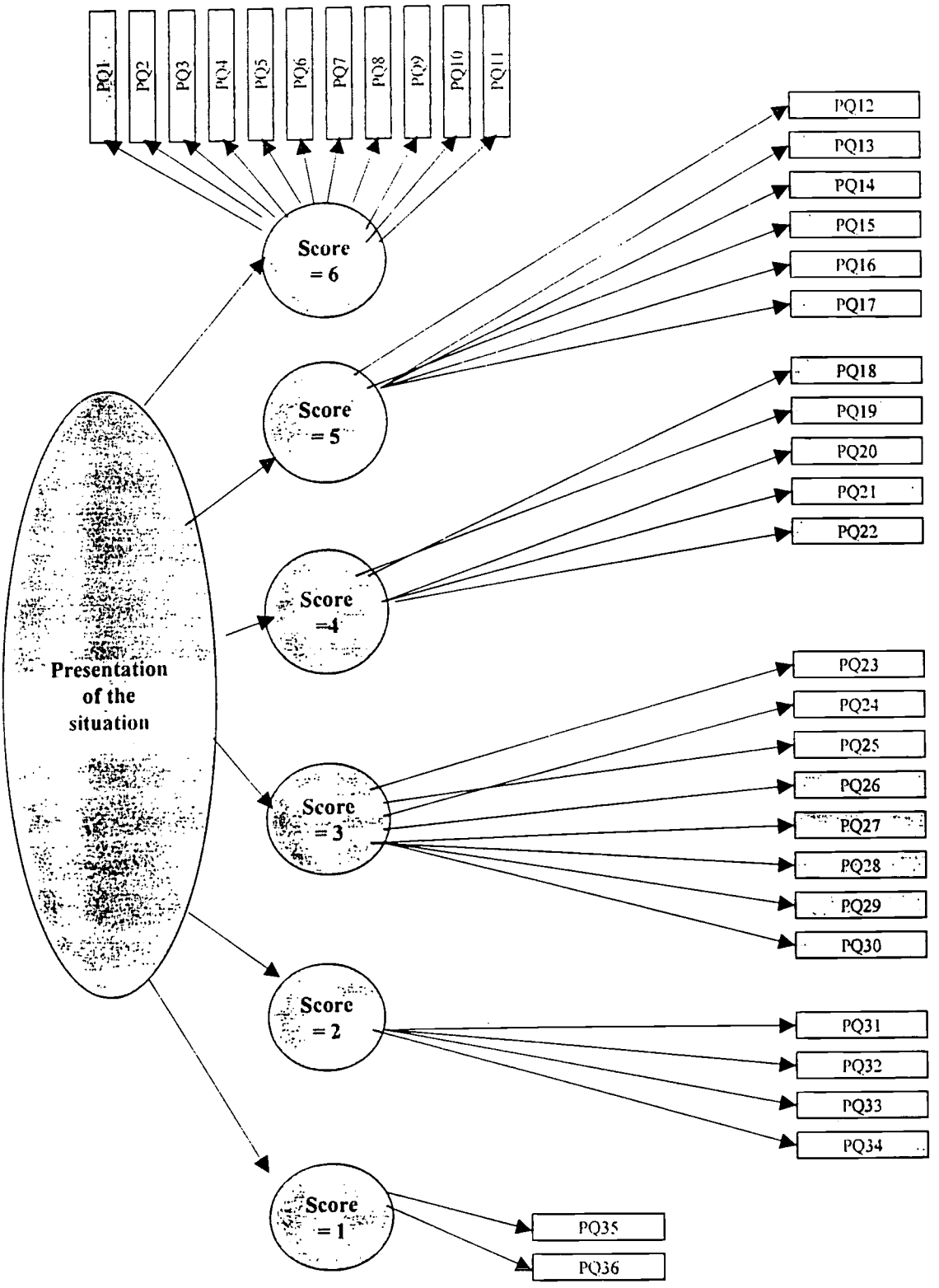


Chart 11

Dimensions indicating the construct of logic and relevance of causes (grade 8)

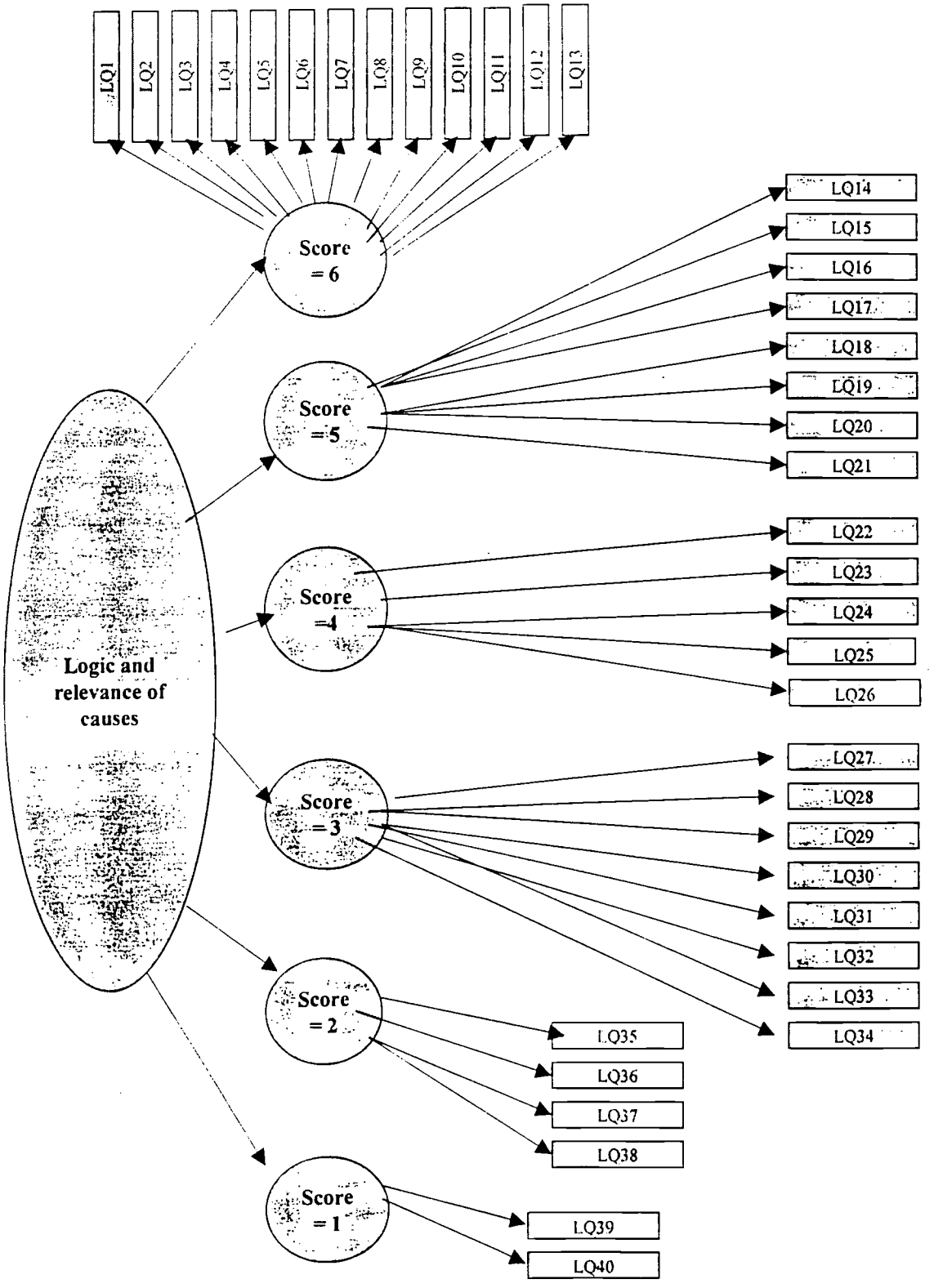
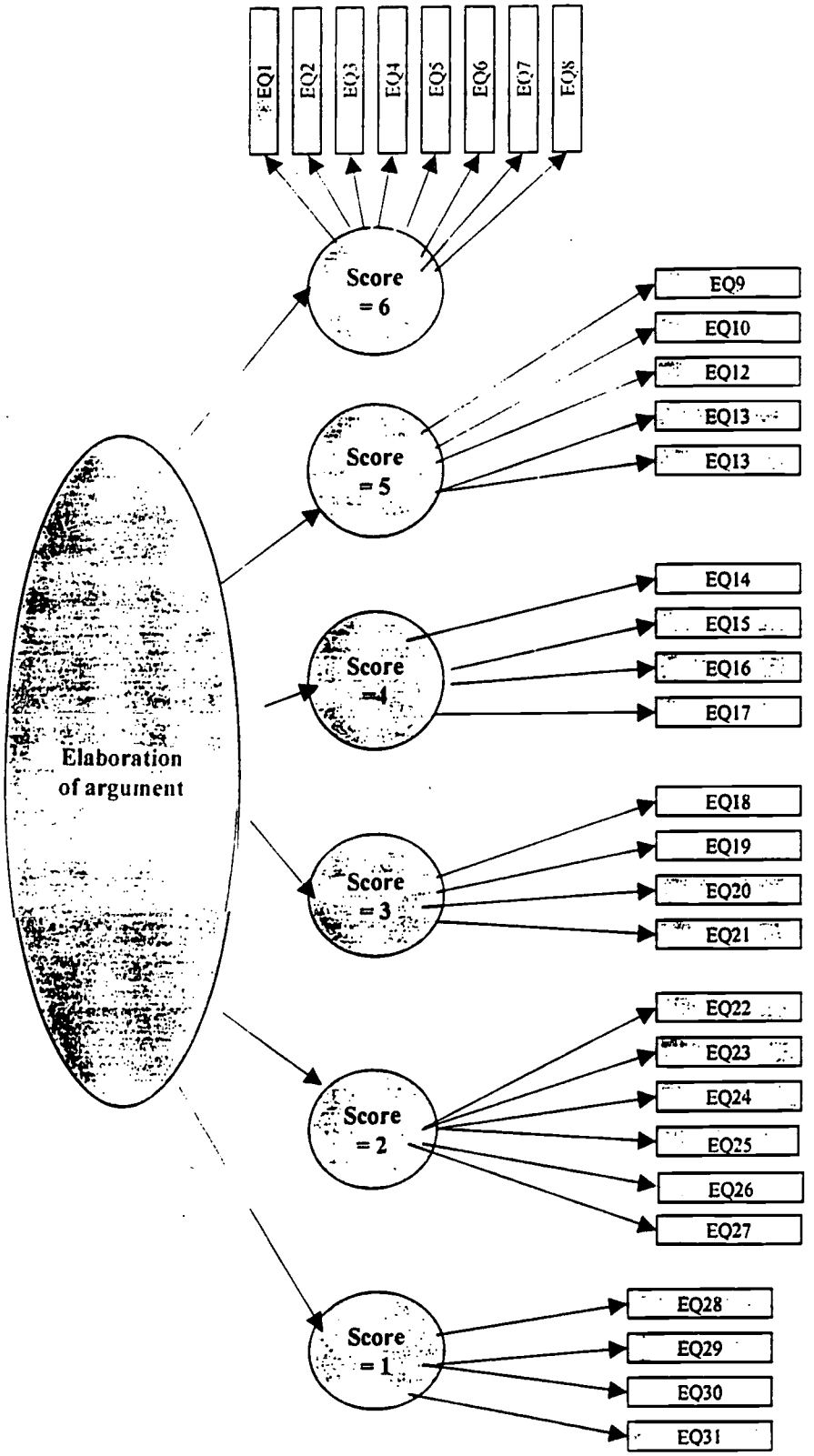


Chart 12

Dimensions indicating the construct of elaboration of argument (grade 8)



WRITING ASSESSMENT

Conventions Scoring Guide
for
ALL TYPES OF WRITING

score points	Criteria
6 **Distinguished Achievement**	**Reader rarely spots errors in conventions** • Usage: Writer demonstrates a thorough command of written English. There may be an occasional lapse due to experimentation with complex ideas and styles. • Mechanics and Spelling: Writer demonstrates command of mechanics and spelling. Minor errors expected in first-draft writing may occur, but they are rare and do not take away from the effectiveness of the writing style.
5 **Noteworthy Achievement**	**Reader seldom spots errors in conventions** **Writer may have difficulty with some more sophisticated conventions** • Usage: Writer demonstrates control of accepted usage; few errors are noticeable • Mechanics and Spelling: Writer demonstrates control of mechanics and spelling but may commit a few errors repeatedly. These do not detract from the overall impression of the essay.
4 **Satisfactory Achievement**	**Reader sometimes spots errors in conventions** • Usage: Writer gives general evidence of an understanding of common usage while committing more than one kind of error. The reader may be aware of these errors but they do not hinder understanding. • Mechanics and Spelling: Writer demonstrates general control of mechanics and spelling but gives evidence of different kinds of errors throughout the essay. The essay may nevertheless be read with relative ease and few distractions.
3 **Some Indication of Achievement**	**Reader finds numerous errors in conventions** • Usage: Writer demonstrates some control of the conventions of usage but frequent errors appear. These usage problems may cause misunderstanding on the part of the reader because the meaning is unclear. • Mechanics and Spelling: Writer shows some control of mechanics and spelling but commits many errors repeatedly. The errors that occur interrupt the flow of ideas, which may confuse the reader.
2 **Limited Indication of Achievement**	**Reader is continually aware of errors in conventions** • Usage: Writer demonstrate limited understanding of the convention of usage. The frequency and seriousness of usage problems may result in a lack of understanding in some areas, but most of the essay can be read with comprehension of the writer's intent. • Mechanics and Spelling: Writer demonstrates little control, committing frequent and serious errors in mechanics and spelling. These errors cause misunderstanding and confusion on the part of the reader.
1 **Few Indications of Achievement**	**Reader is disturbed by repeated errors in conventions** • Usage: Writer demonstrates very little control of conventions of usage, or the piece may be so limited that there is little on which to base a judgment. Essay may be unintelligible. • Mechanics and Spelling: Writer demonstrates little ability in mechanics and spelling. Serious errors may occur with regularity, or the piece may be too brief to evaluate the writer's level of proficiency with English language conventions.

OFF TOPIC papers are scored for conventions only

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM030874

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: A Construct-Centered Generalizability Model: Analyzing Underlying Constructs of Cognitively Complex Performance Assessments	
Author(s): Ying Hong Jiang – Long Beach Unified School District, Philip L. Smith –	
Corporate Source: University of Wisconsin – Milwaukee	Publication Date: April 24, 2000

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1	Level 2A	Level 2B
☒	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature:	Printed Name/Position/Title: Ying Hong Jiang
Organization/Address: Long Beach Unified school District	Telephone: (562) 865-8892 FAX:
	E-Mail Address: yjiang@lbusd.k12.ca.us Date: May 1, 2000

(over)

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

ERIC CLEARINGHOUSE ON ASSESSMENT AND EVALUATION
UNIVERSITY OF MARYLAND
1129 SHRIVER LAB
COLLEGE PARK, MD 20772
ATTN: ACQUISITIONS

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706

Telephone: 301-552-4200
Toll Free: 800-799-3742
FAX: 301-552-4700
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 2/2000)