

M. Ed in Curriculum & Learning program: Learning Technologies (LT) concentration

Reliability checks on LT rubrics – Dec. 2009

Narrative:

Hilary Wilder and Heejung An separately evaluated three samples for each assignment and then met to compare scores. Percentages of agreement are in the charts below. In some cases, the differences could be explained by the fact that one of us was more lenient than the other with a particular student (e.g. sample2 in assessment 8), but after discussion it was apparent that the degree of difference was not that great. The rubrics were tightened up a bit to help with this, but in general, there are some subjective judgments which are going to happen no matter what.

In the case of assignments 6 and 7, after discussion it was decided that the assignment itself needed to be modified so that students would be better able to demonstrate what was expected of them. In assignment 7, another requirement was added (along with another rubric element) which will hopefully encourage the level of synthesis that we are looking for in the assignment. Assignment 6 is a group assignment and as such is a bit trickier since each student's scores is dependent somewhat on their group-mates' efforts. We will be re-thinking this assignment in the future.

Critical Assessment 3: ELCL 612 -- Lesson Plan Database Assignment and Teacher's Technology Skills Checklist Assignment

sample1	80%
sample2	80%
sample3	90%
mean	83%

- After discussion, decided we will use Tech-Integrated Lesson Plan Template to ensure candidates meet all requirements
- Wrote a better defined 'higher-order' for element 5
- Clarified "majority" as ">50%"

Critical Assessment 4: ELCL 611 – Tech Integration Virtual Mentoring Project

sample1	38%
sample2	38%
sample3	38%
mean	38%

- This rubric is to be used in "real-time" as part of the field experience, so after-the-fact reliability checks do not work well.
- Clarified "appropriately" to "as appropriate to the situation" in element 7 & 8

Critical Assessment 6: ELCL 605 -- Social, Ethical, Legal, Etc. Discussions Assignment

sample1	33%
sample2	17%
sample3	67%
mean	39%

- This rubric is for an assignment in which candidates are “topic leaders” for only one discussion, and so, are only expected to meet Target on their assigned topic and meet Adequate on topics that are led by the other candidates in their group. However, if a topic leader is weak, then it is hard for the other candidates in their group to meet the standard. We will explore a backup plan for this situation.
- Since the assignment is to engage in an online discussion, this rubric is also done best in “real-time” and does not lend itself to after-the-fact reliability checks.
- The Safety and Health topic needs to be better defined by the instructor (i.e. it is not cyber safety, but rather physical safety (ergonomics, etc.)).

Critical Assessment 7: ELCL 612 --Technology Skills Rubric and Spreadsheet Assignment Criteria

sample1	25%
sample2	25%
sample3	50%
mean	33%

- A 5th element (paragraph assignment, explaining the areas in need of change based on the aggregated data in the spreadsheet, and giving possible solutions for improving the scores) has been added.
- Clarified “majority” as “>50%”

Critical Assessment 8: ELCL 625 - Technology Grant Proposal Assignment Criteria

sample1	67%
sample2	22%
sample3	89%
mean	59%

- Candidates should be encouraged to get relevant peer-reviewed journal articles for background lit section.
Clarified “majority” as “>50%”
- Clarified expectations for element 3, 5, 6, 8, 9 (e.g. that Overview should provide a clear picture of what will happen during the course of the project and how this will address the objectives of the proposal)

M. Ed in Curriculum & Learning program: Teaching Children Mathematics (TCM) concentration
--

Reliability of TCM Assessments

Two raters, Sandy Alon and Rochelle Kaplan, rated three samples of critical assessments for Assessment 4 (performance in the field/Adapting Instruction for the Inclusive Classroom), Assessment 6 (case study from Math Clinic), Assessment 7 (comprehensive exit requirement/Staff or Parent Development Plan), and Assessment 8 (appreciation of diversity/Equity Plan and Implementation).

These charts indicate the overall percentage of agreement of the two raters for rubric elements and for the assessment as a whole. Below each chart is a brief description of where differences were found and the accommodations that were or will be made in the assessment rubric or course assignment to promote better reliability in ratings.

Assessment 4: On-the-Job Performance – Ideal Lesson Plan with Adaptations, Implementation, and Reflections – Percent of Agreement

Overall	Element 1	Element 2	Element3	Element4	Element5	Element 6	Element 7
16/18 89%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	2/3 66%	2/3 66%

Raters were in agreement regarding 5 of the 7 elements. They agreed on ratings of descriptions of SLOs, assessment of the learning plan, identification of troubleshooting, planned adaptations, and on reflections on the changes for planning in the future based on results.

Differences were found in ratings on the description of the implementation and on reflections of the effect of the lesson on student learning. Since the difference in ratings on the latter two elements was minimal – judging as acceptable vs. target – no adjustments to the rating system were made.

However it was noted that all candidates did poorly in developing a plan for further learning based on implementation results. This seemed to be problem in the assignment which does not seem to cover this category directly. The assignment will be modified to include a separate section on plans for the future.

Assessment 6: Case Study Report on Student from Math Clinic (additional evidence of P-12 student learning and candidates' impacts on student learning and development)-Percent of Agreement

Overall	Element 1	Element 2	Element 3
7/9 78%	3/3 100%	3/3 100%	1/3 33%

Raters were in total agreement on the elements concerning background & behavioral descriptions and descriptions of P-12 students' performance. Disagreement arose on the element concerning the quality of the summary and recommendations. After discussion it was decided that the third rubric element needed clarification and was changed to better reflect the criteria for ratings on this element.

Assessment 7: Comprehensive Exit Requirement – Parent Education or Staff Development Plan (using what was learned throughout the program) – Percent of Agreement

Overall	Element 1	Element 2	Element 3	Element 4	Element 5
14/15 93%	3/3 100%	3/3 100%	2/3 67%	3/3 100%	3/3 100%

Agreement between raters was complete except if the element regarding the description of methods and format of materials. The difference again was slight (acceptable vs. target) and it was decided that no changes were needed in the rubric or in the assignment. However, it was also noted that both raters gave unacceptable ratings to all candidates regarding the section of the assignment concerning the prediction of expected outcomes. It was decided to rework this element of the rubric so that its intention is clearer. It was also decided to add a sentence to the assignment regarding this requirement in order to make the intention clearer to candidates. **(THIS CHANGE WILL BE COMPLETED AT A LATER TIME.)**

Assessment 8: Equity Issues Plan and Implementation Report: Parts I & II (appreciation of diversity is key aspect of assignment) – Percent of Agreement

Overall	Element 1	Element 2	Element 3	Element 4	Element 5
14/15 93%	2/3 67%	3/3 100%	3/3 100%	3/3 100%	3/3 100%

Agreement on this assessment was consistently high. Only on Element 1, regarding the relationship of the identification of an equity issue and a plan to address it, was there any disagreement. It was decided that the rubric would be adjusted to reflect this relationship and in particular focus on the meaning of equity.

Summary of Our Meeting on June 7, 2010

1. Assessment 2 – The title of this assessment will be clearly stated as:

Knowledge of Content and Pedagogy as Evidenced by Grades in Courses Aligned with SPA Standards and Advanced Program Outcomes

The courses included in this assessment beginning in Fall 2010 will be:

ELCL 607, ELCL 608 (the new course on elementary mathematics), ELCL 614, ELCL 615, ELCL 616, ELCL 620, ELCL 626, ELCL 628. (ELCL 613 will be eliminated from this list)

In the Fall, we will revise the curriculum to be consistent with this list and make the following changes.

- 1) Include the 8 ELCL courses listed above as required of all students.
- 2) Write and get approval for the Elementary Mathematics Content course as a permanent course in the program with the alphanumeric, ELCL 608.
- 3) Omit ELCL 619 and TBED 542 as requirement for the TCM concentration
- 4) Have one elective course which may be selected from any other concentration suitable for the candidate's professional work.

2. Assessment 3 - Sandy will revise the rubric for scoring Assessment 3 (part of the ELCL 614 course) – in Fall 2010 - as follows:

Element	Target = 3	Acceptable = 2	Unacceptable = 1
1. Identification of Student Learning Outcomes (SLO's) in plan			
2. Description of forms of representation used in relation to SLO's			
3. Plan for assessment of SLO's			
4. Performance Outcome A during class presentation			

(Sandy to define)			
5. Performance Outcome B during class presentation (Sandy to define)			
6. Performance Outcome C during class presentation (Sandy to define)			
7. Reflections on which SLO's were achieved and which where not			
8. Discussion of how representation formats were or were not successful in the lesson			
9. Discussion of modifications to the lesson for the future based on outcomes noted in #7 and #8			

Criteria for meeting T, A, and U will be adapted from existing rubric, particularly for elements 1, 2, and 3.

3. Assessment 4 – Rochelle modified assignment in December 2009 so that it clearly indicates a section on plans for the future (I suggest renumbering items to conform to scoring elements more clearly as well. To be done in Fall 2010).

4. Assessment 6 – Rochelle modified element 3 on summary and recommendations in December 2009. No further work needed.

5. Assessment 7 – This assessment will be modified in Fall 2010 so that it entails an implementation/performance component. It was suggested that to do this, the plan should involve only one workshop. That workshop will be presented to fellow candidates (and faculty?) either during special sessions set up between 4 and 5 p.m. or during some class sessions. The length of the planned workshop

should be for 1 ½ hours although the actual presentation would not go over 1 hour. The assessment will be scored for planning using the rubric that is currently in place, with a clarification on the element regarding expected learning outcomes from the workshop. An additional rubric for scoring the presentation will be developed. In addition, candidate audience feedback may recorded (modeled on M.Ed. in Reading program practices.)

6. Assessment 8 – Rochelle adjusted the rubric element describing what is meant by the term “equity issue” and plan for increasing equity (December 2009).

7. We will be revising the structure of the program, making a timeline for program completion – especially regarding when the exit requirement needs to be completed, finishing assessment revision and course writing during Fall 2010. It was noted that students will be asked to complete the exit requirement (planning section) prior to beginning Research in Education I.

M. Ed in Curriculum & Learning program: Early Childhood (EC) concentration

Reliability of EC Assessments Fall 2010

In fall 2010, Mary Deblasio and Janis Strasser, rated three samples of critical assessments for Assessment #3 (Lesson Plan). Also in fall 2010, Holly Seplocha and Janis Strasser rated three samples of critical assessments for Assessment #4 (Case Study of Child with Special Needs), Assessment #6 (Advocacy Project), Assessment #7 (Multicultural/Diversity Project) and Assessment #8 (M.Ed. Portfolio).

These charts indicate the overall percentage of the two raters for rubric elements and for the assessment as a whole. Below each chart is a brief description of where differences were found and the accommodations that were or will be made in the assessment rubric or course assignment to promote better reliability in ratings.

The faculty is currently working on redesigning the early childhood concentration M.Ed. program. Critical assessments will be looked at in depth as part of this proposal, which will be submitted to the Curriculum Committee in spring 2011.

Assessment #3 Lesson Plan

Overall	Element 1	Element 2	Element 3	Element 4	Element 5	Element 6	Element 7	Element 8
22/24 91%	2/3 66%	2/3 66%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	3/3 100%

Raters were in agreement on 6 of the 8 elements scored on the rubric. When discussing the differences in scoring for elements 1 and 2, it became clear that these 2 elements are more subjective to personal opinion by raters and each contain more than component of the lesson plan. Element 1 is “content and strategies” and element 2 is “objectives and essential questions.” We will consider redesigning the rubric to separate these components, but we will have to make other changes so that we don’t have more than 8 elements in total.

Assessment #4 Case Study

Overall	Element 1	Element 2	Element 3	Element 4	Element 5	Element 6	Element 7	Element 8
6/6 100%	3/3	3/3						

The raters were in complete agreement when rating the 3 sample written case studies. However, there are 2 other elements which are scored if the student does an oral presentation. The student may choose to do so for 2 of 3 required projects for the course. If she does the oral presentation, she is scored for the content and format of the presentation (Elements 3 and 4). Because Holly Seplocha was not present

for these, she was not able to rate these elements. Elements 1 and 2 rate content and format and are well defined in the rubric.

Assessment #6 Advocacy Project

Overall	Element 1	Element 2	Element 3	Element 4	Element 5	Element 6	Element 7	Element 8
11/12 97%	3/3 100%	2/3 100%	3/3 100%	3/3 100%				

The raters were in agreement in all categories, except for slight difference of opinion in the element “project implementation.” One rater scored “Acceptable” and the other scored “Target.” Perhaps this was due to the wording of the rubric which identified “Project is conducted in a way that advocates for meaningful change” for a Target score and “Project is conducted in a way that advocates for some change” for an Acceptable score. Perhaps these categories need to be more clearly defined.

Assessment #7 Multicultural/Diversity Project

Overall	Element 1	Element 2	Element 3	Element 4	Element 5	Element 6	Element 7	Element 8
6/6 100%	3/3 100%	3/3 100%						

The raters were in complete agreement when rating the 3 sample written Multicultural/Diversity Projects. However, as with the case studies (Assessment #4) there are 2 other elements which are scored if the student does an oral presentation. The student may choose to do so for 2 of 3 required projects for the course. If she does the oral presentation, she is scored for the content and format of the presentation (Elements 3 and 4). Because Holly Seplocha was not present for these, she was not able to rate these elements. However, Elements 1 and 2 rate content and format and are well defined in the rubric.

Assessment #8 M.Ed. Portfolio

Overall	Element 1	Element 2	Element 3	Element 4	Element 5	Element 6	Element 7	Element 8
23/24 98%	3/3 100%	3/3 100%	3/3 100%	2/3 66%	3/3 66%	3/3 100%	3/3 100%	3/3 100%

The rubric is very well defined with each element clearly stated. The raters often score portfolios together at both the graduate and undergraduate level. There was a slight difference of opinion on the philosophy statement, which one rater gave a score of Target, although the student didn’t meet all of the criteria for this element. This was a scoring error.

M. Ed in Curriculum & Learning program: Bilingual/ESL concentration
--

RELIABILITY CHECKS on M.Ed. in Curriculum and Learning, Bilingual/ESL Concentration

Completed February 2011-03-01

Process used for Inter-rater Reliability

Three program faculty members thus far have reviewed program critical assessments. For each assessment, two faculty members conducted a review of three student work samples from each assessment. They used as tools a description of the assignment and the rubric used to evaluate the student outcomes for NCATE purposes for the Bilingual/ESL Concentration of the M.Ed. in Curriculum and Learning and/or the ESL/Bilingual Education Endorsements. Faculty members discussed the evaluations and discussed needed changes to assignments or rubrics when applicable.

Assessments related to M.Ed. thesis and Grades were NOT included in this report.

Data from Inter-Rater Reliability

Our overall Mean Score for agreement on the four assessments listed below is 94.5. Plans to improve reliability are listed under each assessment below.

Assessment #4 –Application of Knowledge as Reflected in Job-Performance: Component of Professionalism in capstone portfolio.

	Professionalism
Sample 1	2/3
Sample 2	3/3
Sample 3	3/3

	% of agreement in Total/Overall Score
Sample 1	66.6%
Sample 2	100%
Sample 3	100%
Mean	88.8

As evidenced above, this assessment produced a high level of reliability. Nonetheless, in our follow-up conversation, we learned that each ranker had a different opinion as to the relative importance of active participation in professional organizations and conferences. In order to solve this disparity, we have agreed to re-do our rubric in line with a similar (but not identical) component of the ESL Endorsement. It

will be necessary to spell out more carefully what degree of participation in organizations can reflect the levels Exceeds, Meets, and Approaches standards.

Assessment 6 – Knowledge of Bilingual Education and/or ESL Field: All sections of capstone professional portfolio except the Section on Professionalism (which is included in Assessment 4).

	Sample 1	Sample 2	Sample 3
Contents	2/3	3/3	3/2
Organization	3/3	3/3	3/3
Personal Reflections	2/2	3/3	3/2
Language	2/2	3/3	3/3
Culture	2/1	3/3	3/2
Instruction	2/2	3/3	3/3
Assessment	2/2	3/3	2/2
Future Directions	2/2	3/3	2/2
Mechanics	2/1	3/2	2/2
Overall Effectiveness	2/2	3/3	2/2

Sample	% of agreement in Total/Overall Score
Sample 1	92.6
Sample 2	98.3
Sample 3	91.1
Mean	94

Overall, this assessment yielded a very high level of reliability.

Some differences came through in the ranking of the CONTENTS page. We decided that new instructions should be given to students.

Also, there was one ranker who felt that two candidates were somewhat lame on cultural theory. We will need to revisit the rubric to determine if more specific explanations of each category is in order.

Assessment 7 – Describing Language: Series of Examinations on English phonetics, phonology, morphology, syntax, pragmatics, and other linguistic components. (ESL Endorsement Assessment 1; Bilingual Education non-SPA report Assessment 2)

	Phonetics/Phonology	Morphology and Syntax	Semantics, Pragmatics Related Concepts
Sample 1	3/3	3/3	2/2
Sample 2	2/2	3/3	2/2
Sample 3	3/3	3/3	2/2

Sample	% of agreement in Total/Overall Score
Sample 1	100
Sample 2	100
Sample 3	100
Mean	100

As demonstrated above, there was 100% agreement between the two raters. This is largely due to the fact that Assessment 7 of the M.Ed. and Assessment 1a of the ESL Endorsement is an objective examination.

Assessment 8 – Diversity Project (Original Term Paper on Multiculturism and/or Acculturation. (ESL Assessment 7; Assessment 2 of non-SPA report for Bilingual Education)

	Theoretical Understanding	Original Theoretical Perspective	Classroom Implications	Integration of Scholarly Sources	Written Communication
Sample 1	2/2	2/2	2/2	2/2	2/2
Sample 2	3/3	3/3	3/3	3/3	3/2
Sample 3	3/3	3/3	3/3	3/2	3/2

Sample	% of agreement in Total/Overall Score
1	100
2	93
3	93
Mean	95.3

This assessment yielded 95.3% of overall agreement between the two readers. Differences were found in the areas of Integration of Scholarly Sources and Written Communication, both which should a 33% difference in rating.

In discussion, the following points were made.

1. One faculty member was going by an older version of APA style, and hence felt the referencing system was incorrect.
2. One faculty member ranked students lower for subjective language (I, in my opinion, the results of my research, etc.)

The first problem can be easily corrected. With regards to the second issue, we have agreed now that only objective language can be used in the assignment.

Summary of Rater Reliability for Assessment 1 and Assessment 5
In the M.Ed. in Curriculum and Learning program

Rater 1 – Dr. Janis Strasser

Rater 2 – Dr. Holly Seplocha

Sample 1 – LB

Sample 2 – BF

2/2011

Both raters agreed on the ratings for Sample 1. We felt it was Target in all areas. For Sample 2, Holly rated the Theoretical Rationale as Acceptable, rather than Target, because it was, in her opinion, too brief (6 pages). Janis felt that, although brief, it met all of the criteria in the rubric to be rated Target. In all other categories, the 2 raters rated Sample 2 the same.

We discussed the formats of the rubrics.. We both felt that the format of the rubrics in awkward to read and could easily be reformatted to combine the rating elements onto the rubric itself. We suggest reformatting the rubrics as shown on the attached sheets.

Assessment 1 - Demonstration of Comprehensive Content and Pedagogical Knowledge
in Field of Specialization-M.Ed. in Curriculum & Learning-all concentrations

Administered in ELCL 630 – Research II- Chapters I and II

RUBRIC

Name of Candidate: _____

Candidate's Concentration: _____

Name of Instructor (Rater): _____

Date/Semester Completed: _____ Score on Assessment: _____

Based on the descriptive criteria, the following is the scoring rubric is suggested for use in evaluating Chapters I and II for comprehensive knowledge of research, theory, and applications in field of specialization.

	1 point- Unacceptable	2 points- Acceptable	3 points - Target	Score
Literature Review	Sparse with references included that are not appropriately related to topic of study or not clearly research-based.	Extensive with most references appropriately related to topic of study within area of candidate's specialization; references are research-based and comprehensively cover the field.	Comprehensively explored with all references appropriately related to topic of study within area of candidate's specialization; major research studies in the field are reviewed.	

Research Questions	No <u>theoretical rationale</u> for study attempted; research questions not posed in answerable formats; variables not defined in assessable terms and/or not consistent with research questions. Demonstrates weak and very incomplete knowledge of major theories in field, trends in field, and standards in field of specialization.	<u>Theoretical rationale</u> for study attempted; research questions posed; variables defined but not fully consistent with research questions or weakly defined in assessable terms. Demonstrates partial knowledge of major theories in field, trends in field, and standards in field of specialization.	<u>Theoretical rationale</u> for study articulated; assessable research questions posed; variables defined and consistent with research questions. Demonstrates full knowledge of major theories in field, trends in field, and standards in field of specialization.	
--------------------	--	---	---	--

Assessment 5 – Evidence of Candidate’s Impact on P-12 Learning
in Field of Specialization-M.Ed. in Curriculum & Learning-all concentrations
Administered in ELCL630 – Research in Education II, Chapters II, IV, V

RUBRIC

Name of Candidate: _____

Candidate’s Concentration: _____

Name of Instructor (Rater): _____

Date/Semester Completed: _____ Score on Assessment: _____

Based on the descriptive criteria, the following is the scoring rubric is suggested for use in evaluating Chapters II, IV and V for comprehensive knowledge of research methods, results, and discussion to determine impact on learning.

	1 point- Unacceptable	2 points- Acceptable	3 points - Target	Score
Research Methods	<u>M</u> inimally revised; may not reflect actual data collection; poorly written with not enough detail to enable replication -methods do not include a plan for assessing P-12 student learning.	Revised and reflects actual data collection; clearly written and enough detail to enable replication - methods include some plan for assessing P-12 student learning.	Revised and reflects actual data collection; clearly written and enough detail to enable replication- methods include clear plan for assessing P-12 student learning.	

Results	Data poorly described and presented; analysis does not reflect answers to posed research questions; no reporting of questions raised and observations during the course of the study - results do not include evidence of impact on P-12 student learning.	Data described and analyzed reflecting answers to posed research questions; minimal or no extension of results to include questions raised and observations made during the course of the study — results include some evidence of impact on P-12 student learning	Data described and analyzed reflecting answers to posed research questions; detailed reporting of questions raised and observations made during the course of the study – results include clear evidence of impact on P-12 student learning.	
Discussions, Conclusions, & Professional Practice Implications	Results not tied to research questions, literature, and relevant theories; no interpretation provided with only restatement of results previously reported; vague generalizations referring to future research and practice in the specialization; recommendations not tied to actual findings of study; recommendations are not tied to specific findings of the study in terms of implications for P-12 student learning and implications for further instruction <i>[Required style formats not used correctly; frequent grammatical errors and typos; Appendices missing; Not acceptable for library]</i>	Results tied to research questions, literature, and theories that were previously articulated, but little interpretation of actual findings and little emphasis on specific implications for future practice in the specialization; recommendations not well tied to specific findings of study; recommendations tied to general findings of the study in terms of implications for P-12 student learning and implications for further instruction <i>[Required style format (e.g., APA) used correctly most of the time; Appendices attached; Not ready for library without substantial revision]</i>	Results tied to research questions, literature, and theories that were previously articulated with logical interpretations of evidence found/reported and strong emphasis on specific implications for future practice in the specialization; recommendations tied to specific findings of the study in terms of implications for P-12 student learning and implications for further instruction <i>[Required style format (e.g., APA) used correctly; Appendices attached; Ready for submission to library or almost ready with correction of a few typos]</i>	

Reliability Data Form on Critical Assessment Ratings

Program Name M.Ed. in Curriculum and Learning Rater's Name Dr. Janis Strasser

Date Completed 2/11

Assessment # 1 Assessment Title Demonstration of Comprehensive Content & Pedagogical Knowledge in Field of Specialization, Administered in ELCL630 Research in Education II, Chapters I and II

RUBRIC SCORES

(Elements)

Sample #	Comments	Element #1	Element #2	Element #3	Element #4	Element #5	Element #6	Element #7	Element #8
1		3	3						
<u>2</u>		3	3						

Scoring: Enter 3 for Target

Enter 2 for Acceptable

Enter 1 for Unacceptable

Reliability Data Form on Critical Assessment Ratings

Program Name M.Ed. in Curriculum and Learning Rater's Name Dr. Holly Seplocha

Date Completed 2/11

Assessment # 1 Assessment Title Demonstration of Comprehensive Content & Pedagogical Knowledge in Field of Specialization, Administered in Research II, Chapters I and II

RUBRIC SCORES

(Elements)

Sample #	Comments	Element #1	Element #2	Element #3	Element #4	Element #5	Element #6	Element #7	Element #8
1		3	3						
<u>2</u>	<u>Ch. 1 too few pges to be rated higher</u>	2	3						

Scoring: Enter 3 for Target

Enter 2 for Acceptable

Enter 1 for Unacceptable

Reliability Data Form on Critical Assessment Ratings

Program Name M.Ed. in Curriculum and learning Rater's Name Dr. Janis Strasser

Date Completed 2/11

Assessment # 5 Assessment Title Evidence of Candidate's Impact of P-12 Learning in Field of Specialization – M.Ed. in Curriculum & Learning – all concentrations, Administered in ELCL630, Research in Education II, Chapters III, IV, V

RUBRIC SCORES

(Elements)

Sample #	Comments	Element #1	Element #2	Element #3	Element #4	Element #5	Element #6	Element #7	Element #8
1		3	3						
<u>2</u>		3	3						

Scoring: Enter 3 for Target

Enter 2 for Acceptable

Enter 1 for Unacceptable

Reliability Data Form on Critical Assessment Ratings

Program Name M.Ed. in Curriculum and learning Rater's Name Dr. Holly SeplochaDate Completed 2/11Assessment # 5 Assessment Title Evidence of Candidate's Impact of P-12 Learning in Field of Specialization – M.Ed. in Curriculum & Learning – all concentrations, Administered in ELCL630, Research in Education II, Chapters III, IV, V

RUBRIC SCORES

(Elements)

Sample #	Comments	Element #1	Element #2	Element #3	Element #4	Element #5	Element #6	Element #7	Element #8
1		3	3						
<u>2</u>		3	3						

Scoring: Enter 3 for Target

Enter 2 for Acceptable

Enter 1 for Unacceptable

M. Ed in Literacy: Reading concentration

Reliability Checks on M. Ed. in Literacy Critical Assessments Rubrics

(Completed Dec 2010)

Process Used for Inter-Rater Reliability

Three program faculty members reviewed program critical assessments. The review included examination of three (3) student work samples from each of the six program assessments, description of assignments, and the rubric used to evaluate the student outcome. The assessment was examined through an independent review by two faculty members who used the rubric to evaluate the three samples. Faculty did not review assessments from courses they taught. After independent review, the group held a discussion to share their evaluations and insights about student outcomes and assignments in relation to program outcomes.

Data from Inter-Rater Reliability Meeting

Overall significant differences were evident in specific rubric elements (see for example assessments 5 and 6) rather than overall scores or grades for candidates. With regard to scores, this inter-rater reliability checks yielded a mean rating above 80% agreement between raters for all assessments. Below are descriptions of specific ratings and comments that emerged from discussion about the outcomes.

Assessment #1: Content Knowledge

Descriptive Title of Assessment: Oral Comprehensive Exam of Master's Thesis

Sample #	Rubric Elements					
	1	2	3	4	5	6
1	4/4	4/4	4/4	3/4	4/4	3/4
	100%	100%	100%	75%	100%	75%
2	4/4	4/4	4/4	4/4	4/4	4/4
	100%	100%	100%	100%	100%	100%
3	3/4	4/4	3/4	3/3	4/4	4/4
	75%	100%	75%	100%	100%	100%

Sample	% of agreement in Total/Overall Score
1	100
2	100
3	91
mean	97

Assessment #2: Content Knowledge – Field Specialization

Descriptive Title of Assessment: Planning and Implementation of a Professional Development Workshop

Sample #	Rubric Elements														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %
2	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %
3	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %	3/3 100 %

Sample	% of agreement in Total/Overall Score
1	100
2	100
3	100
mean	100

Assessment #3: Planning within Specialization

Descriptive Title of Assessment: Action Research Project: Developmental Portfolio on Teaching and Learning

Sample #	Element #1	Element #2	Element #3	Element #4	Element #5	Element #6	Element #7
1	2/3 66%	2/3 66%	2/3 66%	3/3 100%	3/3 100%	3/3 100%	2/3 66%
2	2/2 100%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	2/3 66%
3	2/3 66%	2/3 66%	3/3 100%	3/3 100%	2/3 66%	3/3 100%	3/3 100%

Sample	% of agreement in Total/Overall Score
1	90
2	95
3	86
mean	90

Assessment #4: Application of Knowledge Reflected in Job Performance in the Field

Descriptive Title of Assessment: Diagnosis of Reading Difficulties – Case Study

Sample #	Rubric Elements								
	1	2	3	4	5	6	7	8	9
1	3/3 100%	0/2 0	3/3 100%	2/3 66%	2/3 66%	2/3 66%	2/3 66%	2/2 100%	2/3 66%
2	3/3 100%	0/3 0	3/3 100%	3/3 100%	3/3 100%	3/3 100%	3/3 100%	2/2 66%	2/3 66%
3	3/3	0/3	3/3	3/3	3/3	2/3	3/3	2/3	2/3

	100%	0	100%	100%	100%	66%	100%	66%	66%
--	------	---	------	------	------	-----	------	-----	-----

Sample	% of agreement in Total/Overall Score
1	87
2	92
3	96
mean	92

Assessment #5: Candidate's Impact on Student Learning

Descriptive Title of Assessment: Remediation of Reading Difficulties – Case Study

Sample #	Rubric Elements									
	1	2	3	4	5	6	7	8	9	10
1	2/3	1/3	2/3	2/3	2/3	2/3	2/3	3/3	1/2	3/3
	66%	33%	66%	66%	66%	66%	66%	100%	50%	100%
2	3/3	1/2	3/3	3/3	3/3	2/3	3/3	2/3	1/3	3/3
	100%	50%	100%	100%	100%	66%	100%	66%	33%	100%
3	3/3	2/2	3/3	3/3	3/3	2/3	3/3	3/3	2/2	2/3
	100%	100%	100%	100%	100%	66%	100%	100%	100%	66%

Sample	% of agreement in Total/Overall Score
1	71
2	90
3	93
mean	85

Assessment #6: SPA Specific

Descriptive Title of Assessment: Critical Issues Research Project

Sample #	Element #1	Element #2	Element #3	Element #4
1	2/3 66%	3/3 100%	2/3 66%	2/3 66%
2	2/3 66%	3/3 100%	2/2 100%	3/3 100%
3	1/3 33%	2/3 66%	2/3 66%	3/3 100%

Sample	% of agreement in Total/Overall Score
1	91
2	91
3	67
mean	83

Discussion of Significant Outcomes from Inter-Rater Reliability

1. The assignment for assessment 6 has changed but not the rubric. It was suggested that the rubric be changed so that rubric elements are more aligned to expected outcomes. Review of the samples revealed that some candidates included the technology and multicultural requirements mentioned in the description of the assignment while others did not.
2. It was suggested that the final reflection component of the action research project (assessment 3) become a separate element on the rubric. This rubric element will require that students share what they learned overall. Although this is evident in many of the projects, there is no consistent report of the overall insights gained from the action research project. This rubric element will require an overall report from the teacher regardless of the outcomes observed through the intervention. It was also suggested that a separate element address planning. Some teachers provide specific descriptions of their planning in the teacher reflections section of the

project (element 5), while others include copies of their lesson plans. Changing the rubric will lead to more consistency in teachers' reports about planning. The group also discussed the possibility of making this assessment #7 in the program and extracting the lesson planning elements and outcomes from assessment #5 and using it as assessment #2. This will be discussed further and finalized at the next program meeting where all faculty have the opportunity for input.

3. It was suggested that the rubric elements for assessment #5 should be revised to provide more descriptive information for lesson planning. One possibility discussed in combining elements 2, 3, and 4 into one element for planning. Furthermore rubric elements will list outcomes and descriptions of levels to account for both quality and quantity of lessons taught. It was also suggested that the addition of a scale on the rubric will also help the evaluator rate the candidates' work. The group discussed removing APA for this assessment (element #10) and the possibility of restructuring the rubric to focus each criterion on the skills as expected outcomes.
4. Technology being evaluated through element #8 in assessment # 4 will be rewritten to clarify the kinds of activities that can be used to meet this requirement. The group discussed including a suggested list of activities such as technology-based activities for reading and technology resources for planning will help to define what is meant by technology. The rubric will also identify a specific amount of technology used that would be appropriate for each level of the rubric (e.g. at least 2 technology resources).