



Technical Report

Nebraska State Accountability (NeSA)

Spring 2011 Writing Test

Grades 4, 8, and 11

August 2011



TABLE OF CONTENTS

GENERAL INFORMATION	3
HISTORY.....	3
OVERVIEW.....	4
ADMINISTRATION OF THE WRITING ASSESSMENT	5
WRITING PROMPTS	5
TEST SESSIONS, TIMING, AND FORMAT	5
GRADE 11 NeSA ONLINE PILOT TEST	5
SHIPPING, PACKAGING, AND DELIVERY OF MATERIALS.....	6
MATERIALS RETURN	7
TEST SECURITY MEASURES.....	7
SAMPLE MANUALS	7
PROCESSING AND SCORING THE NeSA-WRITING	8
RECEIPT OF MATERIALS	8
SCANNING OF MATERIALS.....	8
MATERIALS STORAGE	9
PERFORMANCE ASSESSMENT SERVICES (PAS).....	10
GRADES 4 & 8 – HOLISTIC SCORING	10
REPORTING	12
GRADES 4 & 8 REPORTS.....	12
GRADE 11 ONLINE PILOT TEST.....	12
APPENDIX A: NEBRASKA DEPARTMENT OF EDUCATION SCORING GUIDE FOR NARRATIVE WRITING – HOLISTIC – GRADE 4	13
APPENDIX B: NEBRASKA DEPARTMENT OF EDUCATION SCORING GUIDE FOR DESCRIPTIVE WRITING – HOLISTIC – GRADE 8	14

GENERAL INFORMATION

HISTORY

In January 2009, the Nebraska Department of Education (NDE) contracted with Data Recognition Corporation (DRC) to provide and operate a computerized information system to support the administration, record keeping, and reporting for statewide student assessment (NeSA-Reading, NeSA-Mathematics, and NeSA-Science) under the direction of the Department of Education. Legislative Bill (LB) 1157 passed by the 2008 Nebraska Legislature (<http://uniweb.legislature.ne.gov/FloorDocs/Current/PDF/Slip/LB1157.pdf>) requires a single statewide assessment of writing, reading, mathematics, and science in Nebraska's K-12 public schools against the Nebraska academic content standards.

The legislation requires that:

- The assessments will be used for accountability purposes.
- The assessments will be criterion-referenced.

The NDE prescribed such assessments starting in the 2009-2010 school year and phased in as described in Table 1-1. The state uses the expertise and experience of the educators in the state to participate, to the maximum extent possible, in the design and development of the statewide assessment system.

Table 1-1 NeSA Administration Schedule

Subject	Administration Year		Grades
	Field Test	Operational	
Reading	2009	2010	3 through 8 plus 1 high school
Mathematics	2010	2011	3 through 8 plus 1 high school
Science	2011	2012	At least 1 grade in elementary, middle/junior high, and high school

In October 2010 the Nebraska Department of Education (NDE) contracted with Data Recognition Corporation (DRC) to provide and operate a computerized information system to support the administration, record keeping, and reporting for the statewide student NeSA-Writing assessment under the direction of the Department of Education.

NeSA-Writing will be phased in as described in Table 1-2.

Table 1-2 NeSA-Writing Administration Schedule

Year	Paper/Pencil Mode	Online Mode
2011	Grades 4 and 8	Grade 11, optional
2012	Grade 4	Grades 8 and 11
2013	Grade 4	Grades 8 and 11

A governor-appointed Technical Advisory Committee (TAC) consisting of three nationally recognized experts in assessment and measurements, one local administrator, and one teacher from Nebraska provides technical advice, guidance, and research to help NDE make informed decisions regarding standards, assessment, and accountability. There has also been a Statewide Assessment Advisory Group that will continue to provide input into the direction and design of the assessment system from a more local perspective.

OVERVIEW

The NeSA tests are developed specifically for Nebraska. Since 2004, the Nebraska statewide writing assessment has been annually administered in grades 4, 8, and 11 for the purpose of providing school districts with instructional information and to include writing results from grades 4 and 8 as the “other academic indicator” in the federal accountability requirements of the Elementary and Secondary Education Act (ESEA).

The Nebraska statewide writing assessment is intended to:

1. Gather information to assist teachers in determining the progress of students in meeting state or local standards for writing;
2. Provide each local school district with a report of student progress in meeting state or local standards for writing; and
3. Lead to improved writing by Nebraska students.

Data Recognition Corporation (DRC) and Computerized Assessments and Learning (CAL) were the providers of the printed and online versions, respectively, of the 2011 NeSA-Writing Tests.

Paper/Pencil Testing Window: January 24 – February 11, 2011

Online Testing Window: January 31 – February 18, 2011

Number of Potential Testing Sites

254 districts

980 schools

ADMINISTRATION OF THE WRITING ASSESSMENT

WRITING PROMPTS

At each grade level, students responded to a writing prompt developed by NDE to measure composition of writing as specified in the writing content standards. Each student responded to one writing prompt in a specific mode.

Grade 4 – Narrative

Grade 8 – Descriptive

Grade 11 – Persuasive

TEST SESSIONS, TIMING, AND FORMAT

The test window for the grades 4 and 8 paper/pencil tests, including make-up tests, was from January 24-February 11, 2011. The grades 4 and 8 tests were administered in two independent sessions on two consecutive days. Each session was 40 minutes, unless a student’s IEP or 504 Plan required additional time. NDE produced Spanish and large-print versions of these tests for districts that requested them. All student responses were returned to DRC on regular test/answer booklets for processing and scoring.

The required grades 4 and 8 NeSA-Writing paper/pencil tests as well as the voluntary grade 11 NeSA-Writing Online Pilot Test were available to all schools.

Table 2-1 2011 NeSA-Writing Test Participation

Grade	Number of Students Tested Paper/Pencil	Number of Students Tested Online
4	22,234	0
8	21,222	0
11	0	13,972

GRADE 11 NeSA ONLINE PILOT TEST

In 2011, the grade 11 NeSA-Writing Pilot Test was the only assessment administered online. The test window for the pilot test was January 31-February 18, 2011. The purpose of the pilot was to provide an online writing assessment experience for students and schools prior to the testing experience in 2012, so make-ups were not required. The pilot test was administered in one session. Students were allowed to use paper to pre-write and continued their work online by drafting and finalizing their responses. It was recommended by NDE that districts schedule 90 minutes for students to complete the assessment; however, the test was not timed, and students were allowed as much time as necessary to complete and submit their final essays. Spanish and large-print versions were not produced this year as it was a pilot test.

Table 2-2 depicts the percentage of students that completed their online test in each time span.

Table 2-2 2011 NeSA-Writing Grade 11 Online Pilot Test Times

Time Span in Minutes	% in Each Time Span
00-05	0.03%
05-10	0.07%
10-15	0.17%
15-20	0.86%
20-25	1.91%
25-30	3.80%
30-35	6.37%
35-40	8.73%
40-45	10.53%
45-50	11.37%
50-55	10.73%
55-60	10.28%
60-65	8.49%
65-70	6.77%
70-75	5.54%
75-80	4.47%
80-85	3.22%
85-90	2.20%
90+	4.44%

SHIPPING, PACKAGING, AND DELIVERY OF MATERIALS

There was one shipment sent out by Data Recognition Corporation. The shipment was delivered by January 7, 2011. The shipment contained all necessary materials to complete the NeSA-Writing test administration.

- *Writing Manual for Test Coordinators and Administrators*
- Secure Materials: Writing Test/Answer Booklets (Grades 4 & 8)
- Administrative Materials: Student PreID Labels, District/School Labels, Do Not Score Labels, Return Shipping Labels, etc.

DRC ensured that all assessment materials were assembled correctly prior to shipping. DRC Operations staff used the automated Operations Materials Management System (OpsMMS) to assign secure materials to a district at the time of ship out. This system used barcode technology to provide an automated quality check between items requested for and items shipped to each site. A shipment box manifest was produced and placed in each box shipped. DRC Operations staff double-checked all box contents against the manifest prior to the box being sealed for shipment to ensure accurate delivery of materials. Districts and schools were selected at random and examined for correct and complete packaging and labeling.

DRC's materials management system, along with the UPS tracking system, allowed DRC to track the items from the point of shipment from DRC's warehouse facility to receipt at the district. All DRC shipping facilities, materials processing facilities, and storage facilities are secure. Access is restricted by security code. Only DRC inventory control personnel have access to stored secure materials. DRC employees are trained in and made aware of the high level of security that is required.

The assessments for grades 4 and 8 were packaged by school, and shipped to districts to the attention of the District Assessment Contacts. DRC packed 56,470 test/answer booklets, approximately 4,379 manuals, and 4,060 non-secure materials for 1,082 testing sites. DRC used UPS to deliver materials to the testing sites.

MATERIALS RETURN

The materials return window was February 14-18, 2011. DRC used UPS for all return shipments.

TEST SECURITY MEASURES

Test security is essential to obtaining reliable and valid scores for accountability purposes. The 2011 NeSA-Writing included a Test Security Agreement that was provided to all districts by NDE in the *Update: Standards, Assessment, and Accountability (SAA-6) Policies, Practices, Procedures* document. The agreement was to be signed by every school principal and District Assessment Contact and faxed to NDE by January 24, 2011. The purpose of the agreement was to serve as a tool to document that the individuals responsible for administering the assessments both understood and acknowledged the importance of test security. The Test Security Agreement attested that all security measures were followed concerning the handling of secure materials.

SAMPLE MANUALS

Copies of the *Writing Manual for Test Coordinators and Administrators* and the *Online Test Administration Manual* can be found on the Nebraska Department of Education website at www.education.ne.gov/assessment.

PROCESSING AND SCORING THE NeSA-WRITING

RECEIPT OF MATERIALS

Receipt of NeSA-Writing materials began on February 18, 2011, and concluded on March 4, 2011. Any materials received after March 4, 2011, were considered late and were checked-in, scanned, and processed during the late window of March 8, 2011 through May 20, 2011. DRC's Operations Materials Management System (OpsMMS) was utilized to receive secure materials securely, accurately, and efficiently. This system features advanced automation and cutting-edge barcode scanners. Captured data were organized into reports, which provided timely information with respect to suspected missing material.

The check-in process occurred immediately upon receipt of materials; therefore, DRC provided immediate feedback to districts regarding any missing materials based on actual receipts versus expected receipts. DRC produced a Missing Materials Report that listed all test/answer booklets by district, school, and grade that were not returned to DRC.

SCANNING OF MATERIALS

DRC used its image scanning system to capture student essays as images. The images were then loaded into the image scoring system for both the hand scoring of student responses, and for the capture of demographic data.

Customized scanning programs for all scannable documents were prepared to read the writing documents and to electronically format the scanned information. Before materials arrived, all image scanning programs went through a quality review process that included scanning of mock data from production booklets to ensure proper data collection.

After each batch of test/answer booklets was scanned, writing documents were processed through a computer-based edit program to detect potential errors as a result of smudges, multiple marks, and omits in predetermined fields. Marks that did not meet the pre-defined editing standards were routed to editors for resolution.

Before batches of writing responses were extracted for scoring, a final edit was performed to ensure that all requirements for final processing were met. If a batch contained errors, it was flagged for further review before being extracted for scoring and reporting.

MATERIALS STORAGE

Upon completion of processing, student writing test/answer booklets were boxed for security purposes and final storage.

- Project-specific box labels were created containing unique customer and project information, materials type, batch number, pallet/box number, and the number of boxes for a given batch.
- Boxes were stacked on project-specific pallets that were labeled with a list of its contents and delivered to the Materials Distribution Center for final secure storage.
- All paper/pencil test/answer booklets will be securely stored for one year until DRC receives written authorization from NDE requesting that they be permanently destroyed.
- All electronic student response images will be securely stored until DRC receives written authorization from NDE requesting that they be permanently deleted.

PERFORMANCE ASSESSMENT SERVICES (PAS)

GRADES 4 & 8 - HOLISTIC SCORING

Training Material Creation

In 2011, NDE continued use of the established holistic scoring guide, or rubric, for grades 4 and 8. This rubric uses a 1-4 scale to define writing performance holistically. The rubric defines qualities of each score point across six important writing traits.

Four annotated anchor papers from each grade 4 and 8, that were used to illustrate the main score points on the 2010 holistic scoring rubric, were provided to DRC by NDE. In cooperation with NDE, four additional anchor papers for both grades were selected from current field test responses provided to DRC from NDE. These anchor papers, and their justifications, were used to assemble two scoring guides (one for each grade 4 and 8), complete with current training materials provided by NDE. Eventually, two training sets and two qualifying sets were assembled using the field test responses previously scored by NDE.

Reader Recruitment/Qualifications

DRC retains a pool of experienced readers from year to year and many of the 2011 NeSA-Writing readers came from this population. To complete reader staffing for this project, DRC recruiting staff screened applicants for the positions. Candidates were personally interviewed and a mandatory, on-demand writing sample was collected, along with references and proof of a four-year college degree. In this screening process, preference was given to candidates with previous scoring experience, and with degrees in English.

Scoring Directors and Team Leaders were chosen by the content specialists from a pool, consisting of experienced individuals who were successful readers and leaders on previous DRC projects, and who had strong backgrounds in writing. Those selected demonstrated organization, leadership, and management skills. All scoring personnel were required to sign confidentiality agreements before any training or handling of secure materials began.

Training

Representatives from NDE travelled to the DRC Plymouth, Minnesota Scoring Center (February 17-18, 2011) to collaborate with DRC Scoring Directors and Team Leaders during a two-day intense training session.

Two days of reader training took place on February 22 and 23, 2011, at the DRC Scoring Center. Two Scoring Directors, eight Team Leaders, and 78 readers were qualified to score Nebraska grades 4 and 8 student writing responses.

Handscoring Process

Student responses were scored blindly and independently by multiple readers. Readers were not able to see demographic information pertaining to the student being scored, nor were they able to see any of the other scores given by any other reader. Each reader took all traits into consideration when applying a single, holistic score to a given writing response. Each student paper was scored twice and non-adjacent scores were adjudicated.

Data collected from the multiple reads was used to calculate the rater agreement rates and score point distributions.

Quality Control

Validity sets

NDE approved/scored validity responses that were added into the Image Handscoring System for daily quality control checks. These pre-scored responses helped to track how well readers were performing.

Recalibration Tests

During the course of scoring, two recalibration sets were produced using pre-determined scored student responses, and administered to readers as a way to address any scoring issues, and as a method of reinforcing the Nebraska scoring standards set out in the rubric.

Monitoring and Read-Behinds

Team leaders conducted routine read-behinds for every member of their team and provided feedback and assistance to their readers.

Statistical Handscoring Reports

Numerous quality control reports were produced on demand or run daily in order to maintain high standards of scoring accuracy. The inter-rater reliability report and score point distribution report are especially helpful in analyzing scoring data and maintaining high standards of scoring quality.

REPORTING

GRADES 4 & 8 REPORTS

DRC reported student results on the NeSA-Writing for grades 4 and 8. Reports were not printed or shipped to districts/schools. Instead, districts and schools were able to access online reports using DRC's eDIRECT system.

GRADE 11 ONLINE PILOT TEST

The grade 11 online assessment was a pilot of the online system only. No grade 11 student results were reported for the 2011 administration.

NEBRASKA DEPT OF EDUCATION SCORING GUIDE FOR NARRATIVE WRITING										
	1	1+	2-	2	2+	3-	3	3+	4-	4
IDEAS / CONTENT	<ul style="list-style-type: none"> creates no understanding of the events of the story severe digressions from the prompt lacks supporting details storyline is repetitious, disconnected, or seemingly random 		<ul style="list-style-type: none"> creates a limited understanding of the events of the story some digressions from the prompt contains limited, unrelated details storyline is occasionally vague 		<ul style="list-style-type: none"> creates a general understanding of the events of the story is generally focused on the prompt contains adequate, relevant details storyline is generally logical and easy to follow 		<ul style="list-style-type: none"> creates a clear understanding of the events of the story is well-focused on prompt throughout contains numerous, relevant details storyline is distinctive and easy to follow 			
ORGANIZATION	<ul style="list-style-type: none"> structural development does not include a beginning, middle, and end sequencing is random pacing is awkward transitions are missing; connections are unclear 		<ul style="list-style-type: none"> structural development of a beginning, middle, and end is incomplete sequencing is somewhat logical pacing is sometimes inconsistent transitions are predictable, repetitious or weak 		<ul style="list-style-type: none"> structural development includes a functional beginning, middle, and end sequencing is functional and logical pacing is generally controlled transitions are generally effective 		<ul style="list-style-type: none"> structural development includes an effective beginning, middle, and end sequencing is thoughtful, logical and effective pacing is well-controlled transitions clearly show how ideas connect 			
VOICE	<ul style="list-style-type: none"> conveys no sense of the person behind the words tone is not appropriate for the purpose and audience is lifeless and/or mechanical 		<ul style="list-style-type: none"> conveys a limited sense of the person behind the words tone is sometimes not appropriate for purpose and audience is occasionally expressive 		<ul style="list-style-type: none"> conveys a general sense of the person behind the words tone is generally appropriate for purpose and audience is generally individualistic or expressive 		<ul style="list-style-type: none"> conveys a strong sense of the person behind the words tone is well-suited to the purpose and audience is individualistic, expressive, and engaging throughout 			
WORD CHOICE	<ul style="list-style-type: none"> language is neither specific nor precise contains numerous misused or overused words and phrases uses clichés and jargon rather than original language 		<ul style="list-style-type: none"> language is occasionally specific and precise language is occasionally forced or contrived for the purpose and audience few vivid words and phrases some overuse of clichés and jargon 		<ul style="list-style-type: none"> language is usually specific and precise language is generally appropriate for the purpose and audience generally uses vivid words and phrases generally avoids clichés and jargon 		<ul style="list-style-type: none"> language is specific and precise throughout language is natural and appropriate for the purpose and audience effectively uses vivid words and phrases avoids clichés and jargon 			
SENTENCE FLUENCY	<ul style="list-style-type: none"> sentences almost never vary in length or structure choppy, incomplete, rambling, or awkward phrasing throughout fragments or run-ons distract the reader dialogue, if present, is used inappropriately or sounds unnatural 		<ul style="list-style-type: none"> sentences occasionally vary in length or structure phrasing occasionally sounds unnatural fragments, if present, sometimes confuse the reader dialogue, if present, occasionally sounds unnatural 		<ul style="list-style-type: none"> sentences vary generally in length and structure phrasing generally sounds natural and conveys meaning fragments, if present, may add style dialogue, if present, generally sounds natural 		<ul style="list-style-type: none"> sentences vary in length and structure throughout phrasing consistently sounds natural and conveys meaning fragments, if present, add style dialogue, if present, sounds natural 			
CONVENTIONS	<ul style="list-style-type: none"> paragraphing is missing errors in grammar, usage, punctuation, and spelling throughout distract the reader 		<ul style="list-style-type: none"> paragraphing, if attempted, is irregular errors in grammar, usage, punctuation, and spelling may distract the reader 		<ul style="list-style-type: none"> attempts at paragraphing are generally successful a few errors in grammar, usage, punctuation, and spelling—especially with more sophisticated words and concepts—do not distract the reader 		<ul style="list-style-type: none"> paragraphing is sound grammar, usage, spelling and punctuation are generally correct conventions—especially grammar and spelling—may be manipulated for stylistic effect 			

Appendix B: Nebraska Department of Education Scoring Guide for Descriptive Writing – Holistic – GRADE 8

NEBRASKA DEPT OF EDUCATION SCORING GUIDE FOR DESCRIPTIVE WRITING										
	1	1+	2-	2	2+	3-	3	3+	4-	4
IDEAS / CONTENT	<ul style="list-style-type: none"> creates no picture of what is being described severe digressions from the prompt lacks supporting details description is missing 		<ul style="list-style-type: none"> creates a limited picture of what is being described some digressions from the prompt contains some supporting, relevant details description is limited 			<ul style="list-style-type: none"> creates a general picture of what is being described is generally focused on the prompt contains adequate, supporting, relevant details description is acceptable 			<ul style="list-style-type: none"> creates a clear picture of what is being described is well-focused on prompt contains numerous, supporting, relevant details description is distinctive 	
ORGANIZATION	<ul style="list-style-type: none"> structural development does not include an introduction, body, and conclusion sequencing is random pacing is awkward transitions are missing; connections are unclear 		<ul style="list-style-type: none"> structural development of a introduction, body, and conclusion is incomplete sequencing is somewhat logical pacing is sometimes inconsistent transitions may be repetitious, predictable or weak 			<ul style="list-style-type: none"> structural development includes a functional introduction, body, and conclusion sequencing is functional and logical pacing is generally controlled transitions are generally effective 			<ul style="list-style-type: none"> structural development includes an effective introduction, body, and conclusion sequencing is thoughtful, logical and effective pacing is well-controlled transitions clearly show how ideas connect 	
VOICE	<ul style="list-style-type: none"> conveys no sense of the person behind the words tone is not appropriate for purpose and audience is lifeless and/or mechanical 		<ul style="list-style-type: none"> conveys a limited sense of the person behind the words tone is sometimes not appropriate for purpose and audience is occasionally expressive 			<ul style="list-style-type: none"> conveys a general sense of the person behind the words tone is generally appropriate for purpose and audience is generally individualistic or expressive 			<ul style="list-style-type: none"> conveys a strong sense of the person behind the words tone is well-suited to the purpose and audience is individualistic, expressive, and engaging throughout 	
WORD CHOICE	<ul style="list-style-type: none"> language is neither specific nor precise contains numerous misused or repetitious words and phrases lacks vivid words or phrases overuse of clichés and jargon 		<ul style="list-style-type: none"> language is occasionally specific and precise language is occasionally forced or contrived a few vivid words and phrases some overuse of clichés and jargon 			<ul style="list-style-type: none"> language is usually specific and precise language is generally appropriate for the purpose and audience some vivid words and phrases generally avoids clichés and jargon 			<ul style="list-style-type: none"> language is consistently specific and precise throughout language is natural and appropriate for the purpose and audience effectively uses vivid words and phrases avoids clichés and jargon 	
SENTENCE FLUENCY	<ul style="list-style-type: none"> sentences almost never vary in length or structure phrasing is choppy, incomplete, rambling, or awkward fragments or run-ons confuse the reader dialogue, if present, is used inappropriately or sounds unnatural 		<ul style="list-style-type: none"> sentences occasionally vary in length or structure phrasing occasionally sounds unnatural fragments, if present, may confuse the reader dialogue, if present, occasionally sounds unnatural 			<ul style="list-style-type: none"> sentences generally vary in length and structure phrasing generally sounds natural and conveys meaning fragments, if present, may add style dialogue, if present, generally sounds natural 			<ul style="list-style-type: none"> sentences vary in length and structure throughout phrasing consistently sounds natural and conveys meaning fragments, if present, add style dialogue, if present, sounds natural 	
CONVENTIONS	<ul style="list-style-type: none"> paragraphing is missing errors in grammar, usage, punctuation, and spelling throughout distract the reader 		<ul style="list-style-type: none"> paragraphing, if attempted, is irregular errors in grammar, usage, punctuation, and spelling may distract the reader 			<ul style="list-style-type: none"> attempts at paragraphing are generally successful a few errors in grammar, usage, punctuation, and spelling—especially with more sophisticated words and concepts—do not distract the reader 			<ul style="list-style-type: none"> paragraphing is sound grammar, usage, spelling and punctuation are mostly correct conventions—especially grammar and spelling—may be manipulated for stylistic effect 	