

## The New Nebraska Analytic Scoring Model

The NeSA-W scoring model uses one prompt, four domains, and two readers. The domains are *Ideas/Content*, *Organization*, *Voice/Word Choice*, and *Sentence Fluency/Conventions*. Each reader assigns a score to the student's essay for each domain following a rubric developed by Nebraska writing teachers; the domain scores range from a low of 1 to a high of 4. If the scores were simply summed over the domains and readers, the result would be 25 discrete score points ranging from 8 (four 1's from both readers) to 32 (four 4's from both readers.)

### Resolution Scoring

In the unusual situation where the two readers disagree<sup>1</sup> on the score for a domain, the paper is scored on that domain by a third reader. The final score for the domain is then taken to be the two most consistent scores of the three. This will always include the score from the third reader combined with the closest of reader one, reader two, and their average.

### Domain Weights

The composite total score recognizes the relative importance of the four domains by weighting the scores for each. The relative importance of the domains was established as 35%, 25%, 20%, and 20% respectively. The actual calculation is done using weights for the domains of 1.4, 1.0, 0.8, and 0.8, which preserves both the relative importance and the 8-32 range.

### Scale Scores

The *NeSA-W Scale Score* metric was set to range from 1 to 70, with zero reserved for non-scoreable papers. The maximum of 70 was chosen to preserve the significant detail of the possible patterns without exaggerating the precision on the scale. To help understand the Scale Scores, the common score patterns (i.e., 1111, 2222, 3333, and 4444) correspond to Scale Scores of 1, 23, 45, and 70, respectively.

The value of Scale Scores is that they are *measures*, which makes them appropriate for statistical analysis. A change of, say, one Scale Score point has the same significance anywhere along the scale. In contrast, the composite score is just a count of points received and does not reflect the true significance of where the points are assigned. For example, changing 2 to 3 on a paper that was all 2's is substantively much more significant than changing a 2 to 3 on a paper that already was a mixture of 2's and 3's. When a paper is at the "two" level, for example, it has a large impact when a reader breaks the pattern and gives a 3. On the other hand when a paper is between the "two" and "three" levels, it is often difficult for the reader to decide and one more of either has less effect on the estimate of the person's proficiency level. The Scale Score metric recognizes this difference in the person's development.

Calculating the Scale Score is a two step process: (1) computing the composite score by weighting the reader-assigned domain scores and (2) converting the composite score to the Scale Score metric, which requires a computer.

### Composite Scores

The composite score is calculated from the domain scores as:

---

<sup>1</sup> *Disagree* means giving scores that are not adjacent. With this definition, scores of, for example, 2 and 3 are accepted and not sent to resolution; while scores of, say, 2 and 4 will be resolved.

1.

$D_1$ ,  $D_2$ ,  $D_3$ , and  $D_4$  are the sums of the two reader scores for the four domains. The table below illustrates the calculation for a student who received a 2 or 3 from each reader on each domain.

	R1	R2	D=R1+R2	W	W*D
Ideas / Content	3	3	6	1.4	8.4
Organization	2	3	5	1.0	5.0
Voice / Word Choice	3	2	5	0.8	4.0
Sentence Fluency / Conventions	2	2	4	0.8	3.2
Composite Score					20.6

The corresponding Scale Score can then be retrieved from the table below. For the example above, the composite score of 20.6 corresponds to a tabled value for the Scale Score on 35, with a standard error 1.9.

Composite Score	Scale Score	Standard Error
8.0	1	5.0
8.8	7	3.2
9.0	8	2.9
9.4	8	2.6
9.6	9	2.4
9.8	9	2.3
10.0	10	2.2
10.2	10	2.1
10.6	11	2.0
10.8	11	2.0
11.0	11	2.0
11.2	12	1.9
11.4	12	1.9
11.6	12	1.9
12.0	13	1.9
12.2	13	1.9
12.4	13	1.9
12.6	13	1.9
12.8	14	1.9
13.0	14	2.0
13.2	14	2.0
13.4	15	2.0
13.6	15	2.1
13.8	15	2.1
14.0	16	2.2
14.2	16	2.3
14.4	16	2.3

Composite Score	Scale Score	Standard Error
14.6	17	2.4
14.8	17	2.6
15.0	18	2.7
15.2	18	2.9
15.4	19	3.2
15.6	20	3.6
15.8	21	4.2
16.0	23	4.8
16.2	25	4.6
16.4	27	3.9
16.6	27	3.7
16.8	28	3.2
17.0	29	2.8
17.2	29	2.6
17.4	30	2.5
17.6	30	2.4
17.8	31	2.3
18.0	31	2.2
18.2	32	2.1
18.4	32	2.1
18.6	32	2.1
18.8	32	2.0
19.0	33	2.0
19.2	33	2.0
19.4	33	1.9
19.6	33	1.9
19.8	34	1.9

Composite Score	Scale Score	Standard Error
20.0	34	1.9
20.2	34	1.9
20.4	35	1.9
20.6	35	1.9
20.8	35	1.9
21.0	35	2.0
21.2	36	2.0
21.4	36	2.0
21.6	36	2.1
21.8	37	2.1
22.0	37	2.2
22.2	37	2.2
22.4	38	2.3
22.6	38	2.4
22.8	39	2.6
23.0	39	2.7
23.2	40	3.0
23.4	40	3.3
23.6	41	3.8
23.8	43	4.6
24.0	45	5.9
24.2	48	5.9
24.4	50	4.2
24.6	51	3.8
24.8	52	3.3
25.0	53	3.0
25.2	53	2.7
25.4	54	2.5
25.6	54	2.4
25.8	55	2.3
26.0	55	2.2
26.2	55	2.2
26.4	56	2.1
26.6	56	2.1
26.8	56	2.0
27.0	57	2.0
27.2	57	2.0
27.4	57	1.9
27.6	58	1.9
27.8	58	1.9
28.0	58	1.9

Composite Score	Scale Score	Standard Error
28.2	58	1.9
28.4	59	1.9
28.6	59	1.9
28.8	59	1.9
29.0	59	2.0
29.2	60	2.0
29.4	60	2.0
29.6	60	2.1
29.8	61	2.1
30.0	61	2.2
30.2	62	2.3
30.4	62	2.4
30.6	62	2.4
31.0	63	2.9
31.2	64	3.2
32.0	70	5.0