**Emil L. Roy**

# COMPUTERIZED SCORING OF PLACEMENT EXAMS: A VALIDATION

ABSTRACT: *This article validates and refines a computerized system for grading placement exams. To test the reliability of a Structured Decision System (SDS), the author compared computerized ratings with holistic scores, grades earned in writing courses, and ACT-English and ACT Social Studies scores for forty-six placement exams. The study concludes that textual traits linked to levels of writing ability can be quantified, supporting the continued use of well-designed exams to place students in writing courses. Although scores on multiple choice tests generally validate this SDS, they cannot sort levels of writing ability accurately. Further research is needed with a larger population of test-takers, a wider range of test topics, and a greater number of textual traits.*

Hundreds of colleges and schools use timed, written exams to measure the writing ability of incoming students. At their best, these schools limit student misreadings with pretested topics. To achieve uniform responses, many have designed scoring criteria illustrated by carefully selected anchor exams. To achieve consistent holistic rankings by trained readers, they also guard against fatigue, bias, and disagreement. Yet, even at their best, these exams reveal shortcomings. They are expensive to administer and grade, produce an interreader reliability of .90 (Cooper), and provide no feedback to students or instructors beyond unadorned

---

*Emil L. Roy, professor of English at the University of South Carolina-Aiken, has also taught at the University of Southern California, Purdue University, and Northern Illinois University. He is coauthor, along with Sandra Roy, of the* Prentice Hall Guide to Basic Writing, *2nd ed. (Prentice Hall, 1993). His most recent articles were "Direct-Mail Letters: A Computerized Linkage between Style and Success," in* Journal of Business and Technical Communication *(April 1992) and "Evaluating Placement Exams with a Structured Decision System," in* Computers and Composition (April 1992).

© *Journal of Basic Writing,* Vol. 12, No. 2, 1993

numerical ratings. These issues impinge on paper assessment generally, as Pat Belanoff has noted. She cites Ed White's point in *Teaching and Assessing Writing* that "our profession has no agreed upon definition of proficiency and, certainly as a consequence, no agreed upon definitions for proficiencies at various levels of schooling" (58). At the crucial juncture between high school and college writing courses, neither verbal criteria nor the performance of trained readers have approached White's "definition of proficiency." Thus, I have turned to computerized analysis to limit the ambiguities of holistic grading, as applied to impromptu placement exams.

My study used a style checker, RightWriter 4.0 (RW), to analyze forty-six placement exams. A Utah faculty member randomly selected them from more than 2,000 taken. They had been written by students entering the University of Utah in the Fall of 1990. I also used RW to analyze four anchor exams used to illustrate Utah's rhetorical "Criteria," which guide placements by readers. RW measures eight textual traits: readability levels, total number of words, the average numbers of syllables per word and words per sentence, and the percentages of prepositions and unique words. It also creates indexes of "strength" and "descriptiveness."

The Utah Writing Program (Utah) conducts placement testing within a fairly typical set of guidelines. Test-takers have forty-five minutes to write an essay supported by reasons and examples. A prompt asks them to describe a disturbing situation, explain wanted changes, and draw conclusions about people's responses to these situations. Students may consult a dictionary or handbook.

I then designed a Structured Decision System (SDS) with Quattro 2.0's data base capabilities. When based on valid ranking and sorting criteria, my computerized SDS would place papers automatically. Existing research had already established paper length as the most reliable measure of the quality of impromptu writing exams (Brossell, Ruth and Murphy). In a rare computerized study, Reid and Findlay analyzed holistically scored essays with Writer's Workbench. They concluded that longer essays "demonstrate development within paragraphs, structural completeness, and scribal fluency" (12).

My analysis of the four UWP anchor exams singled out two other crucial measurements for identifying especially weak and strong student writing: high syllable averages and low percentages of unique words. The first correlated positively, the second negatively, with fluency. In addition, Reid and Findlay correlated increased word length with essay quality, indicating command of

"a significant vocabulary" and "a mature lexicon" (14). Furthermore, a low percentage of unique words was linked to high writing quality. This confirms injunctions by many writing handbooks to improve organizational completeness and cohesion with repeated words.

Finally, UWP had learned from experience that ratings would apportion themselves predictably. They tend to fall within the four course levels as follows: basic remedial (1-3%), remedial (15-18%), regular composition (60%), and advanced composition (12-16%). My SDS used these parameters to assign placement exams to four acceptance regions; they were divided from one another by precise critical values:

| | Total # of Words | Average # of Syllables | % of Unique Words |
|---|---|---|---|
| 1. Basic Remedial | =<160 | #OR#=<1.2 | #OR#=>66% |
| 2. Remedial | =>161#AND#=<284 | | |
| 3. Regular | =>285#AND#=<495 | | |
| 4. Advanced | =>496 | #AND#=>1.45 | #AND#=<50% |

At this point, the SDS prototype incorporated theoretical underpinnings from reading and rhetorical theory, patterns which emerged from my analysis of four UWP anchor exams, and predictable UWP apportionments of paper ratings.[1]

## Method

Several questions about my SDS's validity remained unanswered:

1. Does the performance of test-takers in writing courses confirm the SDS ratings, establishing their *predictive validity*?
2. Do SDS ratings correlate positively with other tests (such as the ACT-English or Social Studies tests) that measure the same skills, establishing their *concurrent validity*?
3. Does the SDS measure the skills that writing teachers consider important, establishing its *face validity*?
4. Does the SDS measure the essential skills and abilities that comprise the writing competence of professional writers, establishing its *construct validity*?

I based validation of my SDS on several assumptions. First, I predicted that most of the SDS and holistic ratings would match. Further, the remaining divergent rankings of many of the SDS-ranked papers would be borderline; the SDS would rank them no

more than one level higher or lower than their holistic grades.

Thus, the computerized data would verify the face validity of the SDS by passing the "black box" test (Ahituv and Neumann, 506). Apart from links between the SDS criteria and rhetorical theory, the SDS ratings (outputs) would regularly and predictably track the holistic ratings (inputs). I would then need reliable tests to improve the SDS's sorting accuracy of the SDS by adjusting its various critical values up or down.

Second, I also expected to find contradictory SDS and holistic ratings scattered randomly through the fluency rankings. Most, if not all, would involve SDS placements in regular composition where the UWP assigns 60% of the test-takers. To resolve these inconsistencies, I factored in subsequent course grades. For most writing teachers, course grades are more authoritative than holistic ratings. Unlike a timed writing sample, an extended classroom experience provides students with opportunities "to rethink the topic and rewrite their papers, to transform an underdeveloped or incoherent response into a competent essay that will meet the requirements of university discourse outlined in class discussion" (Millward, 109).

Twenty-two of the forty-six takers of placement tests later registered for a writing course in Fall 1990. To resolve rating discrepancies, I assigned significance to course grades. I assumed that a B or C would justify placement in a writing course, whatever its level of difficulty. A grade of D or F indicates that holistic grading placed a student *too high*; the test-taker needed a more elementary writing course. On the other hand, a grade of B+ or A signals that a student's holistic rating was *too low*: the test-taker could have handled a more demanding writing course.

In my experience, nearly all of my B+ or A students have already mastered most of a course's writing skills. They quickly adjust to demands of the course; then they coast, needing a more advanced course to challenge their abilities more fully. Thus, course grades would either confirm the SDS placements, or they would force revisions of critical dividing lines between placement values.

Third, adjustments of the SDS ratings or criteria might still leave unresolved placement discrepancies. To reconcile remaining ambiguities, I would search for anomalies in the RW measurements—extremely high or low counts or indexes. Quite possibly, previously overlooked textual features could override otherwise favorable or unfavorable impressions of an exam.

Fourth, forty-one scores on the ACT-English and Social Studies tests were available for the forty-six samples. I expected these scores, especially on the ACT-English test, to generally validate the revised SDS ratings concurrently. An ACT critical value might even resolve a discrepancy in the ratings. Correlations of these test scores with the SDS ratings might also cast light on the value of multiple choice tests to measure writing ability.

Finally, research which quantifies the textual features of professional writing is still in its infancy, though Garvey and Lindstrom have published a trailblazing study. Thus, construct validity of my SDS cannot yet be verified.

**Results**

Thirty of the forty-six initial SDS and holistic ratings match, for an initial SDS accuracy rate of 65%. Of the sixteen ambiguous ratings, ten are placed within three positions adjacent to the next-rated group. These findings confirm the strong influence of paper length on placement exam ratings. Of the ten borderline SDS rankings, grades in writing courses were available for six. They can correct or refine the critical values dividing one group from another.

My SDS had initially assigned two of the ten exams to basic remedial. Holistic grading placed one of them, Utah18, two levels higher in regular composition. Since the test-taker failed to register for a writing course, no grade exists. However, the exam produced little more than a paragraph (112 words). It also achieved the lowest syllable average (1.2) and highest percentage of unique words (66.96%) in the entire sample. These data confirm the original SDS placement in basic remedial.

The other low-ranked test-taker, Utah44, earned a C in remedial English. This grade validated its holistic placement in this course. It also justified a drop in the SDS fluency floor for remedial placement from =>160 to =>140 words.

In the sample group initially rated remedial (2) by the SDS, three samples were high borderline: Utah5, Utah3, and Utah3 Holistic grading ranked all of them one level higher—in regular freshman composition (3). Two of the three—Utah5 and Utah3—earned high grades in regular freshman composition, an "A" and a "B+," respectively. These grades validated their holistic placements in regular composition. They also warranted a drop in the minimum paper length requirement for this placement from =>284 words to =>243 words.

However, these adjustments raise doubts about the holistic assignment of Utah5's neighbor, Utah9 to remedial writing. The two essays are near-twins, with an identical 243 word count and other similar SDS measurements. Yet, the writer of Utah9 earned a D in remedial writing, confirming its holistic placement. Quite possibly, the test-taker's ACT-English score of 13 signals his or her weaknesses; this was the lowest score for the sample population. In lieu of further research regarding ACT critical values, this score should be added to the SDS remedial algorithm.

The SDS also placed three borderline exams, Utah19, Utah33, and Utah13, in advanced composition, even though holistic grading assigned them to regular composition. Three key measures justify their advanced SDS placement: a high total word count of =>496 added to a high syllable average of =>1.45 *and* a low percentage of unique words: =<50%. In addition, all three test-takers earned A's in their writing courses, validating these SDS critical values for advanced placement.

One remaining borderline exam, Utah40, sends mixed signals. Placed holistically in advanced composition, it falls only one word short of the SDS minimum for this ranking. Its percentage of unique words is commendably low: 43.23%. Yet its use of very short words—averaging 1.3 syllables—would have triggered an SDS placement in remedial writing. However, no grade exists to resolve these discrepancies.

To sum up, of the ten borderline placement exams with divergent SDS and holistic placements, four rankings can be aligned by adjusting the fluency floors for remedial and regular composition downward. One other should be judged a holistic misreading, and course grades confirm three SDS placements in advanced composition. These confirmations and revisions for a total of thirty-eight exams improve the SDS accuracy level to 82.6%. To resolve discrepancies in the ratings for Utah29 and Utah40, some other means is needed.

The SDS initially placed the remaining six exams with divergent rankings in regular composition: Utah20, Utah15, Utah17, Utah37, Utah16, and Utah6. All of them fall well within the SDS acceptance region for regular composition. Yet, holistic grading ranked two of them higher and four of them lower. The low syllable count for three of the six—Utah20, Utah16, and Utah6—justifies their holistic placement in remedial writing. Accordingly, the SDS for this criterion should rise from =<1.2 to =<1.3. In the minds of holistic readers, short, simple words probably overrode the favorable impression created by paper lengths.

At the same time, Utah6 exhibits an interesting anomaly: the lowest percentage of unique words in the sample group: 36.38%. A certain amount of word repetition improves a paper's cohesion. However, trained readers usually prefer mature transitional connections to an overuse of unimaginative repetition. This anomaly suggests that trained readers discriminate against proportions of unique words which are either too high or too low. Perhaps these percentages should not exceed =>66% or drop below =<38%. Adjusting the critical values for the three major SDS criteria resolves no other discrepancies in the ratings.

It seems quite likely, then, that the holistic ratings of a few divergent papers may have been decisively influenced by notable strengths or weaknesses not identified by the dominant SDS measurements. The existence of such anomalies would imply the power of some stylistic traits to sway uncertain holistic graders, toward either the next higher or lower ratings.

Anomalies in their RW measurements may also help resolve discrepancies in the ratings of Utah15, Utah17, and Utah37. Holistic grading placed the first two of them in advanced composition. Yet, on the basis of their modest lengths—327 and 329 words, respectively—the SDS rated them one category lower. While both used commendably long words, their high percentages of unique words suggest writing weaknesses.

However, Utah15 uses the second-longest sentences in the sample population, 23.29 words. No grade is available. Yet the essay's lengthy sentences reinforce its other writing strengths. Reid and Findlay note that "quality impromptu prose . . . does indeed have longer sentences" (13). Tentatively then, sentence lengths averaging =>23.25 words apparently override modest word production, confirming Utah15's holistic assignment to advanced composition.

Utah17 reveals another anomaly, the second highest Flesch-Kincaid readability level in the sample population: 11.44. These readability grades are calculated from a formula combining and weighing average sentence length and average syllables per word. Reid and Findlay note the high correlation of readability scores with essay quality. Like long sentences, a F-K readability level of perhaps =>11.4 should be included in the SDS algorithm for advanced composition. When their numbers are high enough, both sentence length and readability seem capable of overriding mixed ratings in other key stylistic areas.

A third exam with divergent ratings, Utah37, uses the highest percentage of prepositions in the sample population: 15.26%. The

essay's other textual measurements are respectable, even quite good. It uses 367 total words, a 1.52 average syllable count, and a relatively low percentage of unique words: 47.96%. However, the *RightWriter User's Manual* warns that "too many prepositional phrases make the writing wordy and hard to follow" (7-17). Thus, overuse of prepositions confirms the essay's holistic placement in remedial writing. This justifies the addition of a =>15.25% critical value for prepositions to the algorithm for remedial placement.

To sum up, these anomalies apparently exert great influence on the placement of exams which otherwise send mixed stylistic signals. They point to the inclusion of two additional options in the SDS sorting algorithm for *advanced composition*: =>23.25 average words per sentence *or* readability levels of =>11.4. On the other hand, the SDS sorting algorithm for *remedial writing* should include critical values of =>15.26% prepositions *or* =<38% unique words. Like criteria for percentages of unique words, percentages of prepositions probably signify weaknesses in writing if they are either too high or too low. The *RightWriter User's Manual* warns that using too few prepositional phrases "indicates a simple and rigidly structured writing style" (7-16). However, evidence supporting the influence of these critical values is fragmentary. Confirming their effects must await research on a larger population.

None of the data provides any basis for correctly placing the remaining two exams with divergent SDS and holistic ratings, Utah29 and Utah40. Therefore, holistic grading correctly placed thirty-nine of the forty-six samples, for an accuracy rate of 85%. On the other hand, as validated and revised, the SDS places forty-four of the forty-six samples, for an accuracy rate of 95.66%.

The relative influence of RW measurements on SDS ratings can be measured statistically. As indicated by R Squared values, fluency accounts for 42% of the SDS ratings, strength for 21% (a negative correlation), readability for 18%, the average number of syllables for 12%, and words per sentence for nearly 10%. All the other RW variables were less influential.

The ACT-English and Social Studies scores generally validate the SDS rankings concurrently, with one exception. The average ACT-English scores rise from 18.13 for remedial writing to 23.6 for regular composition to 26.4 for advanced composition. The average ACT-Social Studies scores rise from 20.75 for remedial writing to 25.5 for regular composition; however, they drop a point and a half to 24 for advanced composition. This data indicates that a writing sample provides a more reliable basis for distinguishing between levels of writing ability than a highly reputable standardized test.

Finally, the influence of course grades on the computerized rankings of placement exams needs more study. For example, both the SDS and holistic grading place Utah31 in regular composition, where the test-taker earned an A. Does this high grade reflect strong motivation which can't be measured by a placement test? Or does it confirm the predictive powers of a high readability score (11.25)? Two other similarly placed exams, Utah5 and Utah8, also earned A's. Do their grades reflect improvement? Or do their high descriptive indexes, 1.13 and .92, respectively, indicate the predictive force of a neglected RW measurement (involving a writer's use of modifiers)? These questions and others regarding grades cannot yet be answered.

## Conclusions

My SDS for placement tests partly accomplishes Garvey and Lindstrom's goal: "What was once presented as an intuitive benchmark may now be explicable in quantitative terms" (93). This study validates several textual features for rating placement exams:

- Fluency to rank and sort the four rating levels from one another
- Average syllable length and percentages of unique words to sort out basic remedial, remedial, and advanced exam
- Readability level and average sentence length to further distinguish advanced exams
- Percentage of prepositions to sort out remedial exams
- An ACT-English score to identify particularly weak basic remedial exams

Moreover, ACT-English and ACT-Social Studies scores generally correlate with SDS groupings, though (with one exception) they cannot sort placement tests.

A strong case can be made for replacing holistic exam-reading with a still-hypothetical, fully validated SDS. At their best, both systems share a common base: timed impromptu writing samples based on well-designed test prompts, along with an explicit set of grading criteria backed by anchor tests. However, a computerized SDS is superior, in many ways, to holistic grading. It sidesteps interreader disagreements; since it ignores content, exams never trigger reader bias; and no slippage of rating accuracy results from prolonged reader fatigue or boredom.

Basic writing programs, in particular, would benefit from an improved, sharable SDS. No teacher can (or should) say, "For

advanced placement in a writing course, spend no more than forty-five minutes turning out over 500 words averaging at least 1.55 syllables a word and 20 words a sentence. Your writing should also have a readability level of tenth grade or higher, with about 11.5% prepositions and somewhere between 45% and 60% unique words." However, an SDS evaluation system would give in-depth, initial analyses of basic writing students' placement tests to them and their instructors. Equipped with this data, basic writing teachers could more easily arrange teaching priorities. They could also target help with special problems for particular students. Repeated SDS evaluations would also help instructors better measure student progress in crucial areas like scribal fluency, sensitivity to audience, sophisticated use of vocabulary, coherence, and other writing qualities.

Moreover, an SDS evaluation system could help put more electronic classrooms and labs in colleges and more basic writing students in them. This adds up to *empowerment* for basic writing students and instructors alike. Computer-assisted instruction in and outside the classroom provides more accessible answers to the crucial questions: "What's wrong with this writing sample?" and, with the instructor's input, "Why is it flawed?" and "How do I improve it?"

Further study is needed to confirm or correct some SDS critical values, especially those based on isolated anomalies or on course grades. Researchers need to consider the effects on an SDS of differing types of placement topics and prompts (such as description or persuasion). They should also study the influence of other textual features on quality.

A large-scale test of this SDS on a larger population is justified and feasible. However, it is not likely soon. Schools with rigorously designed testing systems are rare; many place daunting barriers between their data and researchers; local and federal funding for such research is virtually nil; the knowledge base is scattered and primitive; and the supply of researchers with the right mix of skills and interests is extremely limited, even nonexistent. When one adds the staunch resistance of what one reader terms "sec-hume Luddites" in the profession, movement toward a sharable, validated SDS for placement exams is bound to be slow. Yet, as many writing teachers already know, the computer is a new and authoritative tool, with many underutilized and undiscovered applications to the writing situation.

University of Utah Placement Exams

| Code Name | Reada- bility | Total #Words | Ave# Sylbls | #Wrds Sent | %Wrds Preps | St'gth | Desc't | %Uniq Words |
|---|---|---|---|---|---|---|---|---|
| Sort Basis: | | | | Basic Remedial [1] | | | | |
| =<139#OR#=<1.2 | | | | | | | | #OR#=>66% |
| Utah18 | 7.94 | 112 | 1.5 | 13.90 | 13.39% | 0.84 | 0.40 | 66.96% |
| Sort Basis: | | | | Regular Remedial [2] | | | | |
| #OR#=>140 #OR#=<1.21#OR#=>15.25% | | | | | | | | #OR#=<38% |
| #AND#=<242#AND#=<1.3 | | | | | | | | |
| Code Name | Reada- bility | Total #Words | Ave# Sylbls | #Wrds Sent | %Wrds Preps | St'gth | Desc't | %Uniq Words |
| Utah44 | 8.73 | 150 | 1.6 | 13.55 | 8.00% | 0.57 | 0.86 | 50.67% |
| Utah24 | 10.68 | 167 | 1.46 | 23.14 | 13.17% | 0.00 | 0.57 | 65.27% |
| Utah45 | 5.37 | 173 | 1.37 | 12.28 | 12.72% | 0.62 | 0.86 | 59.54% |
| Utah11 | 9.87 | 176 | 1.58 | 17.50 | 11.90% | 0.33 | 0.64 | 61.93% |
| Utah14 | 7.40 | 178 | 1.46 | 14.75 | 8.99% | 0.62 | 0.71 | 63.48% |
| Utah9 | 7.99 | 243 | 1.39 | 18.39 | 12.76% | 0.36 | 0.65 | 49.79% |
| Utah20 | 4.58 | 302 | 1.3 | 12.50 | 10.93% | 0.69 | 0.73 | 47.68% |
| Utah16 | 8.18 | 368 | 1.27 | 22.63 | 8.40% | 0.06 | 0.73 | 45.38% |
| Utah6 | 8.70 | 437 | 1.3 | 22.95 | 8.90% | 0.00 | 0.73 | 36.38% |
| COUNT | | 9 | | | | | | |
| % OF TOTAL | | 19.57% | | | | | | |
| MIN | 4.58 | 150 | 1.27 | 12.28 | 8.00% | 0 | 0.57 | 36.38% |
| MAX | 10.68 | 437 | 1.6 | 23.14 | 13.17% | 0.69 | 0.86 | 65.27% |
| AVG | 7.94 | 243.8 | 1.41 | 17.52 | 10.64% | 0.36 | 0.72 | 53.35% |
| ST/DV | 1.85 | 97.1 | 0.11 | 4.27 | 1.96% | 0.27 | 0.09 | 9.19% |

Sort Basis: Regular Composition [3]
=>243
#AND#=<494

| Code Name | Reada-bility | Total # Words | Ave # Sylbls | # Wrds Sent | %Wrds Preps | St'gth | Desc't | %Uniq Words |
|---|---|---|---|---|---|---|---|---|
| Utah5 | 8.18 | 243 | 1.5 | 15.13 | 10.29% | 0.30 | 1.13 | 61.32% |
| Utah3 | 8.16 | 259 | 1.5 | 14.88 | 11.58% | 0.71 | 0.57 | 52.12% |
| Utah41 | 5.45 | 277 | 1.35 | 13.09 | 10.83% | 0.70 | 0.61 | 56.32% |
| Utah23 | 8.46 | 290 | 1.44 | 18.06 | 8.28% | 0.30 | 0.77 | 49.31% |
| Utah1 | 9.59 | 297 | 1.44 | 21.07 | 12.12% | 0.14 | 0.62 | 53.54% |
| Utah39 | 8.74 | 297 | 1.52 | 16.44 | 13.47% | 0.43 | 0.70 | 52.86% |
| Utah21 | 9.12 | 309 | 1.5 | 18.10 | 10.36% | 0.32 | 0.78 | 52.43% |
| Utah35 | 9.71 | 337 | 1.57 | 17.47 | 13.95% | 0.35 | 0.63 | 49.85% |
| Utah2 | 10.14 | 343 | 1.59 | 18.00 | 12.83% | 0.37 | 0.70 | 51.90% |
| Utah42 | 3.97 | 347 | 1.35 | 9.35 | 9.79% | 0.69 | 0.71 | 49.86% |
| Utah46 | 6.16 | 350 | 1.34 | 15.13 | 11.43% | 0.47 | 0.65 | 56.57% |
| Utah30 | 10.92 | 364 | 1.58 | 20.17 | 13.18% | 0.38 | 0.71 | 54.95% |
| Utah31 | 11.25 | 365 | 1.64 | 19.16 | 11.00% | 0.17 | 0.95 | 47.95% |
| Utah37 | 11.13 | 367 | 1.52 | 22.50 | 15.26% | 0.00 | 0.73 | 47.96% |
| Utah22 | 5.55 | 379 | 1.4 | 13.50 | 11.87% | 0.69 | 0.68 | 51.72% |
| Utah27 | 7.18 | 385 | 1.49 | 13.21 | 11.43% | 0.54 | 0.52 | 55.58% |
| Utah7 | 8.60 | 423 | 1.35 | 21.10 | 11.35% | 0.00 | 0.76 | 47.52% |
| Utah4 | 7.53 | 432 | 1.28 | 20.48 | 9.49% | 0.21 | 0.75 | 44.91% |
| Utah32 | 7.15 | 451 | 1.42 | 15.52 | 9.50% | 0.42 | 0.62 | 49.22% |
| Utah38 | 7.66 | 452 | 1.48 | 15.03 | 11.28% | 0.34 | 0.71 | 40.93% |
| Utah43 | 8.32 | 460 | 1.49 | 16.39 | 10.22% | 0.47 | 0.80 | 46.96% |
| Utah12 | 9.85 | 462 | 1.4 | 21.95 | 10.17% | 0.13 | 0.92 | 39.83% |
| Utah25 | 9.37 | 468 | 1.48 | 17.21 | 10.47% | 0.32 | 0.66 | 50.43% |
| Utah8 | 8.54 | 481 | 1.33 | 21.73 | 13.93% | 0.32 | 0.92 | 45.11% |
| **Utah29** | **6.25** | **493** | **1.37** | **15** | **11.77%** | **0.51** | **0.72** | **40.16%** |
| **Utah40** | **7.38** | **495** | **1.3** | **19.60** | **11.52%** | **0.20** | **0.63** | **43.23%** |
| Utah10 | 7.66 | 506 | 1.38 | 18.00 | 11.07% | 0.39 | 0.79 | 42.49% |
| COUNT | | 27 | | | | | | |
| % OF TOTAL | | 58.70% | | | | | | |
| MIN | 3.97 | 243 | 1.28 | 9.35 | 8.28% | 0 | 0.52 | 39.83% |
| MAX | 11.25 | 506 | 1.64 | 22.5 | 15.26% | 0.71 | 1.13 | 61.32% |
| AVG | 8.22 | 382.7 | 1.44 | 17.31 | 11.42% | 0.37 | 0.73 | 49.45% |
| ST/DEV | 1.75 | 78.02 | 0.09 | 3.16 | 1.55% | 0.19 | 0.13 | 5.29% |

Sort Basis: Advanced Composition [4]

#AND#1.45

#OR#=>11.4#OR#=>495#OR#=>23.25 #AND#=>50%

#AND#=<65%

| Code Name | Reada-bility | Total # Words | Ave # Sylbls | #Wrds Sent | %Wrds Preps | St'gth | Desc't | %Uniq Words |
|---|---|---|---|---|---|---|---|---|
| Utah15 | 11.21 | 327 | 1.56 | 23.29 | 7.95% | 0.00 | 0.73 | 58.41% |
| Utah17 | 11.44 | 329 | 1.57 | 21.86 | 9.12% | 0.01 | 0.92 | 52.89% |
| Utah19 | 8.91 | 496 | 1.5 | 17.70 | 12.70% | 0.20 | 0.61 | 46.98% |
| Utah33 | 7.58 | 499 | 1.45 | 15.56 | 10.60% | 0.32 | 0.66 | 45.89% |
| Utah13 | 12.84 | 499 | 1.63 | 23.70 | 14.80% | 0.10 | 0.64 | 49.70% |
| Utah26 | 8.67 | 518 | 1.45 | 18.46 | 11.39% | 0.15 | 0.78 | 46.72% |
| Utah28 | 9.59 | 532 | 1.56 | 17.47 | 12.40% | 0.23 | 0.70 | 46.62% |
| Utah36 | 10.14 | 537 | 1.59 | 17.80 | 11.17% | 0.13 | 0.81 | 47.49% |
| Utah34 | 11.39 | 576 | 1.53 | 23.00 | 12.85% | 0.10 | 0.82 | 45.31% |
| COUNT | | 9 | | | | | | |
| % OF TOTAL | | 19.57% | | | | | | |
| MIN | 7.58 | 327 | 1.45 | 15.56 | 7.95% | 0 | 0.61 | 45.31% |
| MAX | 12.84 | 576 | 1.63 | 23.7 | 14.80% | 0.32 | 0.92 | 58.41% |
| AVG | 10.20 | 479.2 | 1.54 | 19.87 | 11.44% | 0.14 | 0.74 | 48.89% |
| ST/DEV | 1.57 | 84.2 | 0.06 | 2.90 | 1.95% | 0.10 | 0.09 | 4.01% |
| R Sq Values for SDS Regression | 0.185 | 0.41 | 0.126 | 0.095 | 0.008 | 0.217 | 0.085 | 0.06 |

## Note

[1]This study entitled "Evaluating Placement Exams as a Structured Decision System" was published in *Computers and Composition* 9.2 (Apr. 1992): 71-83.

## Works Cited


Ahituv, Niv, and Seev Neumann. *Principles of Information Systems for Management.* 3rd ed. Dubuque, IA: Wm. C. Brown, 1990.

Belanoff, Pat. "The Myths of Assessment." *Journal of Basic Writing* 10.1 (1991): 54-66.

Brossell, Gordon. "Current Research and Unanswered Questions in Writing Assessment." *Writing Assessment: Issues and Strategies.* Ed. Karen L. Greenberg, Harvey S. Wiener, and Richard A. Donovan. New York: Longman, 1986. 168-82.

Cooper, Charles. "Holistic Evaluation of Writing." Ed. Charles R. Cooper and Lee Odell. *Evaluating Writing: Describing, Measuring, Judging.* Urbana, IL: NCTE, 1977.

Garvey, James J., and David H. Lindstrom. "Pros' Prose Meets *Writer's Workbench*: Analysis of Typical Models for First-Year Writing Courses." *Computers and Composition* 6.2 (1989): 81-109.

Millward, Jody. "Placement and Pedagogy: UC Santa Barbara's Preparatory Program." *Journal of Basic Writing* 9.2 (1990): 99-112.

Reid, Stephen and Gilbert Findlay. "*Writer's Workbench* Analysis of Holistically Scored Essays." *Computers and Composition* 3.2 (1986): 6-32.

*RightWriter User's Manual.* Sarasota, FL: Que, 1990.

Ruth, Leo, and Sandra Murphy. *Designing Writing Tasks for the Assessment of Writing.* Norwood, NJ: Ablex, 1988.