AN INVESTIGATION OF THE IMPACT OF A TRAITS-BASED WRITING MODEL ON STUDENT ACHIEVEMENT

PAPERWORK REDUCTION ACT CLEARANCE REQUEST

SUPPORTING STATEMENT PART B

Prepared For:

Institute of Education Sciences
United States Department of Education
Contract No. ED-06-CO-0016

Prepared By:

Northwest Regional Educational Laboratory
Center for Research, Evaluation, and Assessment
March 2007

Table of Contents

List of Exhibits

SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION

B. DESCRIPTION OF STATISTICAL METHODS

1. Respondent Universe and Sampling Methods

The study will not statistically sample students or schools from a fixed universe, but will identify and recruit eligible schools in Oregon containing 5th grade classrooms and a minimum of 30 5th grade students. Eligible schools will be those that are not already thoroughly implementing a 6-trait based writing approach, and that are willing to participate in the research protocol. The sample will consist of schools that are willing and able to participate and that are not already implementing the intervention. There are approximately 555 schools with 5th grade classrooms and a 5th grade enrollment of 30 or more; it is unknown how many of these are not already thoroughly implementing a 6-trait based writing approach, or how many of these will be interested in participating.

Accordingly, the universe of cases will be those 5th grade students in schools which volunteer to participate in the study, and other students/schools like them. Since the schools are not randomly sampled from a known population, generalization to other schools cannot be made strictly on statistical grounds based on this single study. In other words, it cannot be known with certainty whether any other specific school or student is part of the population from which this sample was drawn. Rather, as in most group randomized trials, generalization to other sites must be based on considered judgments about the similarity of those sites to schools in the study, and/or schools from replication studies in additional sites and under additional circumstances.

In a cluster randomized design, the response of units of randomization is the key response rate issue. In this study, schools are the unit of randomization. Response rates for schools are expected to be at or very near 100% since participating schools will have chosen to participate.

2. Procedures for Collection of Information

<u>Statistical Methodology For Stratification and Sample Selection</u>

From the pool of eligible elementary schools, 64 schools will be selected and half will be randomly assigned to the treatment group that will implement the traits-writing model, and half will be assigned to the control group (for a total of 32 experimental schools and 32 control schools). Random assignment will be completed independently by Chesapeake Research Associates to ensure that such assignment is independent of school recruitment efforts.

Eligible schools will be those that are not already thoroughly implementing a 6-trait based writing approach, and that are willing to participate in the research protocol. Oregon uses a trait-based framework for statewide writing assessments in grades 4, 7 and 10, but the state provides little support for integration of this framework in classroom teaching and learning. Some schools provide teachers with professional development in trait-based writing instruction to fill this gap. A screening process will exclude those schools in which elementary level teachers are already thoroughly trained in this model and are already integrating this approach in their classroom instruction and student assignments. It is expected that principals and teachers at schools which have not made strong efforts in the past to use this approach in their classrooms will be interested in participating in the study, in order to receive the professional development that is provided. Furthermore, these schools will provide a reasonable counterfactual, with teachers who have some awareness of the popular trait-based writing models but who have not been thoroughly trained and have not used this model with students.

Randomization of schools will occur within strata defined by districts. Within districts, the two schools with the highest percentage of students eligible for free or reduced price lunch will be randomly assigned to conditions, then the next highest pair of schools, and so on, so that the experimental and control groups will be reasonably well balanced in terms of socio-economic status (SES) and variables that correlate highly with SES. In districts with an odd number of participating schools, the unpaired school will be randomly assigned to condition.

Within a participating school, all 5th grade writing teachers will participate regardless of previous writing content skills, as will al of their 5th grade students (there will be no sampling within schools). Typical Oregon schools have two 5th grade teachers which will yield approximately 64 teachers and 1,536 students in each condition at baseline, i.e., 64 teachers and an average of 24 students per class. Assuming that as many as one sixth of students present at baseline may not complete the post-test, yielding a conservative estimate of 20 students per classroom, the final student sample for this analysis is expected to be at least 1,280 in the experimental group and 1,280 in the control group.

Student attrition, as well as crossovers and students who enter study classrooms after baseline, will be monitored. Attrition rates will be reported in detail; if attrition is large, or different across conditions, the final sample may be weighted during the analysis to preserve the representative nature of the sample. Students who cross over from treatment schools to baseline schools will be included in the treatment sample for the main analysis, which will be an intent-to-treat (ITT) analysis. If particular teachers within the treatment group fail to implement the model in their classrooms (“no-shows”), those teachers and students will still be included in the treatment group for the ITT analysis. Students entering study schools during the year will not be included in the analysis.

Data Collection Plans

Data collection will occur for two groups of schools corresponding to the 2007-08 and 2008-09 school years. Activities for each group will be coordinated for three data levels: student, teacher, and school.

Student achievement in writing will be assessed with holistic and trait scores of a written essay from all participating students in September and May of the experimental year. The writing essays will be administered by participating teachers from both control and treatment groups. This is consistent with the Oregon statewide writing assessment. NWREL staff has developed a student assessment booklet that will include the essay prompt and instructions to guide the student through a three-day process of planning, drafting, and finalizing an essay. Teachers will also provide coded, non-identifiable data on a few demographic and assessment-related variables for each student using the Coded Student Data Form. Teachers will return completed forms to Chesapeake Research Associates (CRA) in a stamped, addressed envelope, where they will be recorded and prepared for scoring by NWREL raters.

The teacher survey will be completed by all teachers in both treatment and control groups prior to the beginning of treatment (September), in mid-year (February), and again toward the end of the treatment year (May). NWREL will provide teachers a copy of the survey and a stamped, addressed return envelope. NWREL will record receipt of surveys and follow-up via e-mail or telephone with teachers who do not return the survey within two weeks. School-level data, including demographic and student achievement data, will be collected in the summer preceding the treatment year via the internet.

Estimation Procedures

The impact of the intervention on student performance will be estimated using an intent-to-treat (ITT) analysis, in which data from all subjects will be analyzed as part of the group to which those subjects were originally randomly assigned. We believe that crossovers and no-shows will be minimal, therefore no provisions are included below for separate estimates of impact on the treated (IOT). If this assumption proves to be flawed, the plan may be revised to include supplemental analyses bracketing the lower and upper bounds for the IOT estimate.

Student writing scores will be the dependent variable, predicted primarily by membership in the treatment or control group. Several covariates will be used to increase the efficiency of the design. All students with post-test scores will be included in the analyses; baseline scores will be imputed for any students who do not complete the pre-test. A hierarchical linear model will be used to analyze the treatment and control group differences in student achievement. The data analysis will be accomplished using a mixed model ANOVA in which the effect of the experimental manipulation is estimated as a fixed effect, while the effects of school level variables and the individual differences among teachers and students will be estimated as random effects. Specific data analysis techniques are described below for each research question:

Research Question 1: What is the impact of 6+1 Trait® Writing on student achievement in writing?

The analysis of student outcomes will have two components. First, outcome data will be subjected to a series of descriptive analyses to examine the distributions and insure that statistical assumptions are reasonably met for the primary analyses. Second, outcome data will be subjected to a series of inferential analyses for hypothesis testing, which will involve the comparison of the treatment group and the control group. Since our data will have a nested structure, we will utilize HLM (hierarchical linear modeling) for this purpose. This will be done for the entire data set, then repeated for the subgroups. Subgroup analyses, however, can also be carried out by entering the subgroup identity as a fixed factor at Level-1 (student-level).

In the following, we will outline the statistical models that will be used for hypothesis testing. Multilevel modeling techniques will be used to properly account for the nested structure of the data, i.e., the fact that observations within groups are not independent. The purpose of the multilevel modeling is, therefore, primarily to obtain an efficient, unbiased estimate of the impact of the treatment. Covariates will be used primarily to improve the efficiency of the model by controlling extraneous variance, although some covariates will also be explored as possible moderators of the treatment effect.

Two models are presented in the following. The first model is a 2-level CRT with Level 2 Covariate. The second model is the generalized 2-level CRT, which can have numbers of covariates at both levels. This model is presented for the purpose of discussing the possible expansion of the model.

(1) Two-level CRT with Level 2 Covariate

- Level 1 Model (i.e., Student Level Model)

$$Y_{ij} = \beta_{0j} + e_{ij} \qquad e_{ij} \sim N(0, \sigma^2)$$

- Level 2 Model (i.e., School Level Model)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + \gamma_{02}S_j + u_{0j} \qquad u_{0j} \sim N(0, \tau_{|s})$$

where:

- y_{ij}: outcome measure of student i at school j.
- β_{0j}: mean outcome measure of students at school j.
- e_{ij}: residual associated with each student. It is assumed to be normally distributed with the mean of 0 and the variance of σ^2.
- γ_{00} : grand mean for the outcome measure.
- γ_{01}: treatment effect.
- W_j: indicator variable. Treatment group is indicated by 0.5; control group, -0.5.
- γ_{02}: coefficient for the school-level covariate.
- S_j: school-level covariate, which is the school mean for the previous year's 5th grade test score.
- u_{0j} : residual associated with the school mean of the outcome measure. It is assumed to be normally distributed with the mean of 0 and the variance of $\tau_{|s}$.

In this model, the test score of a student is defined as the school-level mean plus the random error associated with each student. The school-level mean, in turn, is defined as the grand mean plus the effect of the treatment plus the random effect associated with the school. The school-level mean, however, is adjusted for the covariate which is the previous year's school mean.

The random assignment to conditions will take place at the school level in the study. Consequently, the treatment effect will show up at the school level. Reducing the error variance at the school level, therefore, will result in the gain in the power.

The HLM model above is essentially a 2-level nested mixed-model ANCOVA, in which the treatment and the school level covariate are entered as fixed effect variables, whereas the cluster (school) is entered as the random effect variable.

(2) Generalized Two-level CRT

The model can be expanded by identifying and entering more covariates. Such covariates can be either at the student or school level. The following shows the general two-level HLM with multiple covariates.

- Level 1 Model (i.e., Student Level Model)

$$Y_{ij} = \beta_{0j} + \beta_{1j}a_{1ij} + \beta_{2j}a_{2ij} + \ldots + \beta_{pj}a_{pij} + e_{ij}$$

- Level 2 Model (i.e., School Level Model)

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_{1j} + \gamma_{02}W_{2j} + \ldots + \gamma_{0s}W_{sj} + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}W_{1j} + \gamma_{12}W_{2j} + \ldots + \gamma_{1s}W_{sj} + u_{1j}$$

:

:

$$\beta_{pj} = \gamma_{p0} + \gamma_{p1}W_{1j} + \gamma_{p2}W_{2j} + \ldots + \gamma_{ps}W_{sj} + u_{pj}$$

Theoretically, any number of covariates could be entered. However, more covariates means that a larger sample will be necessary to estimate the coefficients. Given the limitation in the sample size we can afford, especially in the context of experimental studies in which the effects of covariates are randomly distributed across conditions, our plan is to use the simplest model that reflects the data structure accurately enough. This means exercising due diligence in choosing only good covariates, with a maximum of four covariates.

Research Question 2: How do student impacts vary by pre-existing characteristics of schools, teachers and students?

Additional analyses will be performed to provide further information about the impact of the 6+1 Trait® Writing model by uncovering factors that moderate the level of impact of the treatment. The primary concerns here are that student or school variables examined must not be reactive to treatment condition, and a valid counterfactual must be available and identifiable in the control condition.

Analyses of demographic subgroups will be performed in this context, as well as the contribution of a measure of prior exposure of teachers to trait-based writing models. These additional factors will be entered at appropriate levels into the model specified in the primary analysis. For instance, the measure of prior exposure will be entered into the Level-2 (teacher level) equation. Subgroup analyses, though exploratory, will also be performed by entering such school level variables as "%[subgroup]" into the Level-3 equations, while entering a student level variable "[subgroup] membership" into the Level-1 equation. These analyses are intended for aiding interpretation of the result of the primary analysis.

Each step in the above data analyses will be fully documented and reported, including the processes and techniques used and the detailed results and effect sizes. Student level demographic variables that may have implications when aggregated at the classroom or school level will initially be modeled at each level, and removed from particular levels if no effects or interactions are found. For example, race/ethnicity may have effects at the level of individual students, but may also have classroom or school effects when aggregated to those levels. Additional exploratory analyses will be conducted to examine relationships between level of implementation and impact on student achievement.

Statistical Power Estimates

To determine the level of statistical power attainable from various sample sizes, a set of analyses was performed to estimate the minimum detectable effect size under a set of scenarios. The minimum detectable effect size was defined here as the necessary size of effect in order to maintain statistical power of 0.8. The power analysis presented in the following was specifically performed for the detection of the main effect of treatment on student outcome.

Collaborating with the Oregon Department of Education, we were able to calculate the ICC for a 2-level CRT model very similar to that proposed for the current study (student nested within School). The ICC was calculated using the 2005 – 2006 Grade 4 Oregon Writing Assessment data (N=39057). First, the unconditional ICC was calculated by fitting a 2-level CRT model without any covariate. Then, the conditional ICC was calculated by fitting a 2-level CRT model with the previous year's building mean as the school-level (L2, for Level 2) covariate. The effect of this covariate, R^2_{L2}, was calculated from the school-level variance in the two models. Specifically:

- When a 2-level CRT model without covariate was fit to the data, school-level variance (τ) was 2.917, and student-level variance (σ^2) was 18.221. The unconditional ICC was, therefore, calculated as $\tau / (\tau + \sigma^2) = 0.138$.

- When a 2-level CRT model with the covariate was fit to the data, school-level conditional variance ($\tau_{|x}$) was 1.574, and student-level variance (σ^2) was 18.225. The conditional ICC was, therefore, calculated as $\tau_{|x} / (\tau_{|x} + \sigma^2) = 0.079$.

The effect of the covariate, R^2_{L2}, was calculated from the initial school-level variance (τ) and subsequent the school-level variance conditional to the use of covariate ($\tau_{|x}$).
$R^2_{L2} = 1 – (\tau_{|x} / \tau) = 0.428$. It appears that the previous year's school mean (school-level covariance) and the true school mean are moderately to strongly correlated, $R_{L2} = 0.654$.

Based on this information, a power analysis was performed for the present study using the Optimal Design Software, the unconditional ICC of 0.14, and the effect of the covariate (R^2_{L2}) of 0.43. Exhibit 1 shows the number of schools required to attain MDEs ranging from .10 to .25.

Exhibit 1. Statistical Power Estimates

Unconditional ICC	R^2_{L2}	Number of Schools	Minimum Detectable Effect Size (at power of .8)
0.14	0.43	320	.100
		144	.150
		84	.200
		54	.250

The general goal of the power analysis was to estimate the necessary number of schools to sample, in order to maintain power of 0.8 for a minimum detectable effect size of $\delta = 0.25$. In performing the power analysis, the number of students per teacher (n) was set to 20. The number of teachers per school (J) was set to 2. The default value of 0.05 was used for the alpha level.

The results of this power analysis indicate that the planned sample size of 64 elementary schools will provide adequate power to detect an MDE of .24 or larger given no school level attrition, and enough power to detect an MDE of .25 or larger even with considerable school-level attrition, which we hope to avoid. The number of schools included in Wave 2 of the study can be adjusted up or down depending on the findings of preliminary analyses of data collected from the Wave 1 schools. Actual power for the study is likely to be somewhat better than the conservative estimate presented above, since it does not take into account the individual level baseline writing measure as a covariate. The degree to which an individual level covariate helps with power at the school level is unpredictable and isn't modeled in the Optimal Design software.

Unusual Problems Requiring Specialized Sampling

There are no such unusual circumstances.

Use of Periodic Data Collection Cycles to Reduce Burden

This is a one-time research study.

3. Methods to Maximize Response Rates and to Deal with Issues of Non-response

We expect very high response rates for this study. Since the study is conducted within a single school year, as part of schoolwide curriculum and instruction that is implemented for all students in the 5th grade classrooms, it is unlikely that attrition or non-response of students or teachers will be great or that it will differ across conditions. Some students may be absent for one of the assessments, and some students will likely move into or out of each school. This will be monitored closely. Study personnel will have regular contact with participating teachers and will have ample opportunities to facilitate the collection of teacher surveys and student essays. As noted above, the key issue here is not non-response of individuals, but non-response of entire schools, since schools are the unit of assignment and analysis, and school estimates are robust to the minor expectable loss of individual responses. Recruitment of schools will include procedures to insure that the school as a whole is willing to complete the study.

4. Pilot Testing of Instruments

Pilot testing is expected for the Teacher Survey. Only nine or fewer respondents will be involved for the pilot testing. Teachers will be asked to complete each survey as if they were in the full-scale study. They will be asked to confirm their time to complete and will be debriefed in a phone or face-to-face interview about their experiences with the questionnaire to ensure that items are clear and that they are gathering the information intended.

5. Contractor Name Responsible for Design, Analysis, and Data Collection for the Study

This study will be conducted by the Center for Research, Evaluation and Assessment, the Northwest Regional Educational Laboratory (NWREL), under the Regional Educational Laboratory Contract with the Institute of Education Sciences, the U.S. Department of Education. Chesapeake Research Associates (CRA) is providing consultation services around research design and analysis for the study.

Michael Coe	Principal Investigator	NWREL	503-275-9497
Gary Nave	Project Analyst	NWREL	503-275-9573
Makoto Hanita	Project Analyst	NWREL	503-275-9628
Michael Puma	Project Consultant	CRA	410-897-4968
David Connell	Project Consultant	CRA	410-897-4968