Abstract Title Page
Not included in page count.

Title:
Instructional Variability in Bilingual Education Programs: Time of Year, Raters, and Content

Authors and Affiliations:
Lee Branum-Martin. Texas Institute for Measurement, Evaluation, and Statistics, University of Houston. Lee.Branum-Martin@times.uh.edu

Paras D. Mehta. Texas Institute for Measurement, Evaluation, and Statistics, University of Houston. Paras.Mehta@times.uh.edu

Coleen D. Carlson. Texas Institute for Measurement, Evaluation, and Statistics, University of Houston. Coleen.Carlson@times.uh.edu

David J. Francis. Texas Institute for Measurement, Evaluation, and Statistics, University of Houston. dfrancis@uh.edu

Barbara R. Foorman. Florida Center for Reading Research, Florida State University. bfoorman@fcrr.org

Abstract Body
Limit 4 pages single-spaced.

Background / Context:
        There are many types of programs for Spanish speaking students in the US, with varying methods and goals (Baker, 2001; García, 1997; Tabors & Snow, 2001). Some preliminary work suggests that bilingual classrooms may differ widely in instruction, even under the same program labels (Branum-Martin, Foorman, Francis, & Mehta, 2010; Branum-Martin et al., 2006; Branum-Martin et al., 2009; Cirino, Pollard-Durodola, Foorman, Carlson, & Francis, 2007; Foorman, Goldenberg, Carlson, Saunders, & Pollard-Durodola, 2004; Saunders, Foorman, & Carlson, 2006). However, there are few studies which have compared the extent to which various bilingual program models differ in actual instruction delivered.

Purpose / Objective / Research Question / Focus of Study:
        Directly measuring instructional practice however, is difficult and costly, involving the influence of time, raters, content, and programs (Raudenbush, 2008).The purpose of the current paper is to estimate the relative influence of these important sources of variance in classroom observations completed in a large quasi-experiment of bilingual education.

Setting:
        Schools were selected from Texas and California, representing urban areas in both states. In addition, schools were selected from the Texas border region near Mexico.

Population / Participants / Subjects:
        Thirty two schools were selected which met acceptable academic performance criteria in their respective state, had 40% or more Hispanic students, and used one of three educational programs: English immersion, dual language, and transitional education. The observations were completed on 315 teachers (85% female, 75% Hispanic, 20% White), by 27 trained observers with experience in bilingual education.

Intervention / Program / Practice:
        The classroom observation instrument used in the study (Foorman, et al., 2004; Foorman & Schatschneider, 2003) was adapted from Scanlon and Vellutino (1996) to quantify time spent on various reading/language arts behaviors and to include language used during instruction. Using a tape-recorded designation of minutes, observers coded the content of teaching, and teacher language use on a minute-by-minute basis. All observations were conducted by trained project research assistants. Training involved the review, explanation, and discussion of all the codes, coding practice based on videotaped lessons, and live coding practice in classrooms with reliability checks conducted by site coordinators. Only those who achieved acceptable levels of reliability during practice sessions were allowed to conduct formal classroom observations (see Foorman et al., 2004, for descriptions of content codes, training, and reliability).
        The protocol included a total of 24 content codes which were summed into two instructional domains. The first domain, Oral Language, includes: oral language/discussion, listening comprehension, language strategies, and vocabulary. The second domain, Reading and Language Arts (RLA), included: book and print awareness, discussion of predictable text, phonemic awareness, alphabetic instruction, structural analysis, word work, spelling, reading text (teacher reads aloud, students read aloud, students read silently), writing composition, and

grammar/ capitalization/ punctuation/ mechanics. Non-instructional time (breaks, transitions, and interruptions) was also coded, but is not analyzed here.

Research Design:

  These observations came from a larger longitudinal quasi-experiment designed to follow Spanish speaking children from kindergarten through second grade, sampled from classroom programs of instruction. The observations were made three times per year. In the three years of the project across the three grades, 924 observations were completed, with an additional 122 (11%) observations with an additional rater present in the room to check reliability ($r > .80$; Foorman et al., 2004). The counts of the 924 completed observations on the 315 teachers are shown in Table 1.

<center>(Table 1 here)</center>

Data Collection and Analysis:

  In order to estimate variability due to repeated measures within teacher, rater, and school, we fit versions of the following level 1 model for time i, teacher j, in school k:

$$Y_{ijk} = \pi_{0\,ijk} + \pi_{1\,ijk} * \text{time}_i + \pi_{2\,ijk} * \text{Program}_{\bullet jk} + \pi_{3\,ijk} * time_i * Program_{\bullet jk} + e_{ijk}$$

where Y_{ijk} represents the observed number of minutes in one form of instruction at time i for teacher j in school k, $\pi_{1\,ijk}$ represents the effect of time (linear or dummy-coded by wave), $\pi_{2\,ijk}$ represents the effect of program (dual or immersion versus transitional), and e_{ijk} represents measurement error (assumed to be normally and independently distributed across observations).

  For simplicity, we present the level 2 (teacher) and 3 (school) equations together, with random effects specified in italicized words (non-Greek):

$\pi_{0\,ijk} = \gamma_{000} + teacher_{0\,\bullet jk} + rater_{a\,\bullet jk} + school_{0\,\bullet\bullet k}$	[intercept]
$\pi_{1\,ijk} = \gamma_{100} + teacher_{1\,\bullet jk}$	[time slope]
$\pi_{2\,ijk} = \gamma_{200}$	[program effect]
$\pi_{3\,ijk} = \gamma_{300}$	[time*program]

where γ_{000} represents the grand intercept, with random effects for teacher, rater (primary a, and secondary b), and school. The second equation, $\pi_{1\,ijk}$, represents the overall effect of time, γ_{100}, plus a teacher-specific linear slope, $teacher_{1\,\bullet jk}$. The final equation represents the effect of program, with an effect for overall mean difference, γ_{200}, and school-specific random deviation, $school_{2\,\bullet jk}$.

  In this way, we partition the variability in each domain of observation (RLA and oral language instruction for Spanish and English) into components representing time, teacher, rater, and school. In this model, teacher and rater represent cross-classified random effects. Significant variability across raters represents systematic differences in how the observers perceived instructional actions. With 32 schools, we do not consider school level variability around the main program effect (γ_{200}), and linear slope (γ_{100}).

Findings / Results:

  Table 2 shows the means and SD for each grade, program, time point, language, and domain. The means show the potential for large differences across programs. We fit models with a linear effect of time for the 3 observations per teacher. We also fit models without the linear

trend (a dummy code for semester) and found little substantive differences. We also graphed teacher trend lines and did not detect strong evidence of a nonlinear group trend or clusters of teachers with severe nonlinear trends. We therefore take variability in the linear slope as a crude indicator of time variability at the teacher level.

(Table 2 here)

Table 3 presents the results for the model of English Reading observations. The top portion of the table presents fixed effects and the bottom presents random effects. There are three columns, one for each grade, kindergarten through second. For kindergarten, the intercept of 11.7 indicates that for a completely average teacher and rater in Fall in a transitional education classroom, the model-predicted average was 11.7 minutes of English reading and language arts instruction. There was no average linear change (-0.5 min) across semesters. The dual program had a large (-10.0 min) but imprecise (not significant) difference compared to the transitional program. However, teachers in English immersion classrooms on average taught 22.9 minutes more English reading than teachers in transitional program classrooms. There was a slight program by time interaction in which dual program teachers taught 8.7 minutes per semester more than transitional teachers.

(Tables 3-6 here)

The fixed effects for the other grades show a similar pattern, in that there is no substantial linear change, no difference between dual and transitional programs, and immersion programs tend to teach more English than transitional programs.

The random effects at the bottom of Table 3 show estimates of the variability due to the factors in the model. The school intercept variability of 166.9 yields a SD of 12.9 minutes in the average amount of English reading instruction between schools. The SD for teachers was 5.2 minutes, and the SD for teachers' linear change was 4.8 minutes per semester—neither was statistically significant. The school, teacher intercept, and teacher slope variability accounted for 45%, 7%, and 6% of the total variability, respectively. Rater effects had an SD of 4.5 minutes (5% of the overall variance). The residual SD was 11.4 minutes.

The random effects for the other grades in Table 3 show moderate school variability (12% to 13%) and high teacher variability in intercept (48 to 66%), and a fair amount in teacher slope (4% to 6%). Rater effects were small, ranging from 0% to 6%.

Table 4 shows results for Spanish Reading instruction. In the fixed effects, there was only a significant linear trend in first grade, and the immersion program teachers used significantly less time in Spanish instruction. Because of the few teachers and schools and small variability in instruction, many random effects could not be estimated. Rater effects were too small to be estimated, except in kindergarten. Teachers, however, varied greatly, by 14.4 to 21.5 minutes, accounting for 34% to 62% of the variance among observations. Teacher variability appeared to increase across the grades.

Table 5 shows results for English oral language instruction. In the fixed effects, there were no average differences across programs in English oral language instruction, except in kindergarten. There was no linear change and no teacher variability in change. Teachers were somewhat variable (SD = 1.6 to 3.7 minutes). Rater variance was small or unable to be estimated. Schools differed from each other essentially as much as teachers differed from each other. Overall, English oral language instruction appeared fairly homogeneous.

Table 6 shows results for Spanish oral language instruction. In the fixed effects, there was no significant linear trend, except in kindergarten. Teachers in English immersion spent on average less time in Spanish oral language instruction, except in second grade. In the random

effects, school differences were small (1.5 to 3.0 minutes SD). While the percentage of variability across teachers appeared high (19% to 53%), it that was small in magnitude (SD = 2.2 to 5.5 minutes).

Conclusions:

The results suggest sharp differences across the instructional domains and languages. English reading and language arts yielded the most stable models in terms of estimated random effects. The restricted models in the other domains may imply that there is less variability in those aspects of this design in this sample. The intraclass correlations should be interpreted carefully, especially in the oral language domains, where variability was low and the percentages appear high.

The results for English reading and language arts instruction suggest little average change from fall to spring within years. However, there was substantial variability around these effects. Teachers differed in their fall to spring rates of change, implying that some were changing instruction as the year progressed. These large differences across teachers imply that even within program, actual delivered reading instruction can vary greatly.

This study is limited by the choice of two instructional domains per language: reading and oral language (Saunders, et al., 2006). Other methods which allow for more simultaneous content codes may be informative, but the 24 content categories and the dependence of choosing one code over the other pose tough analytic challenges (a large, sparse multinomial model). This study has applied a 3-level cross-classified model to the total minutes in two domains, but bivariate Spanish-English and cross-domain models will be informative next steps. In addition, the additional 11% reliability ratings not analyzed here represent a multiple membership model, which may provide additional guidance on rater effects.

This study is also limited by having only 3 time points per year. More time points could allow for more sensitive exploration of the nature of within-year change. A final limitation we leave for the next step in our work is to examine these estimates of instruction in relation to student performance.

Overall, English immersion differed from transitional instruction most sharply in reading instruction in the expected manner: more English, less Spanish. There were far fewer differences in oral language instruction (either English or Spanish), and across grades, program differences appeared to decrease. This implies that the major instructional difference between the programs may lie more in reading instruction and less so in oral language instruction, at least in these early grades. The results may imply that both immersion and primary language programs (dual and transitional) converge in decreasing the use of Spanish oral language instruction by second grade.

The lack of strong rater differences appears to support the training protocol. The lack of linear change or variability in change either implies that teachers are highly consistent over time, or that more frequent observations are needed to effectively index within-year change (there was variability in instructional change, at least in English reading). It is not clear that more raters are needed, but perhaps a design with more time points could reveal stronger evidence on the need for more frequent observations.

These measures show high reliability and temporal stability and may serve as better indicators of instruction than the program labels. We look forward to using estimates based on this approach in models of language and literacy achievement among students, classrooms, and schools.

Appendices
Not included in page count.

Appendix A. References
References are to be in APA version 6 format.

Baker, C. (2001). *Foundations of bilingual education and bilingualism* (3rd ed. Vol. 27). Clevedon, England: Multilingual Matters.

Branum-Martin, L., Foorman, B. R., Francis, D. J., & Mehta, P. D. (2010). Contextual effects of bilingual programs on beginning reading. *Journal of Educational Psychology, 102*(2), 341-355.

Branum-Martin, L., Mehta, P. D., Fletcher, J. M., Carlson, C. D., Ortiz, A., Carlo, M. S., & Francis, D. J. (2006). Bilingual phonological awareness: Multilevel construct validation among Spanish-speaking kindergarteners in transitional bilingual education classrooms. *Journal of Educational Psychology, 98*(1), 170-181.

Branum-Martin, L., Mehta, P. D., Francis, D. J., Foorman, B. R., Cirino, P. T., Miller, J. F., & Iglesias, A. (2009). Pictures and words: Spanish and English vocabulary in classrooms. *Journal of Educational Psychology, 101*(4), 897-911.

Cirino, P. T., Pollard-Durodola, S. D., Foorman, B. M., Carlson, C. D., & Francis, D. J. (2007). Teacher characteristics, classroom instruction, and student literacy and language outcomes in bilingual kindergartners *Elementary School Journal, 107*(4), 341-364.

Foorman, B. R., Goldenberg, C., Carlson, C. D., Saunders, W., & Pollard-Durodola, S. D. (2004). How teachers allocate time during literacy instruction in primary-grade English language learner classrooms. In P. McCardle & V. Chhabra (Eds.), *The voice of evidence in reading research* (pp. 289-328). Baltimore, MD: Brookes.

Foorman, B. R., & Schatschneider, C. (2003). Measurement of teaching practices during reading/language arts instruction and its relationship to student achievement. In S. Vaughn (Ed.), *Reading in the classroom: Systems for observing teaching and learning*. Baltimore, MD: Brookes Publishing.

García, O. (1997). Bilingual education. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 405-420). Oxford, UK: Blackwell Publishers.

Raudenbush, S. W. (2008). *Statistical inference when classroom causality is measured with error*. Paper presented at the Society for Research on Educational Effectiveness, Washington, DC.

Saunders, W. M., Foorman, B. R., & Carlson, C. D. (2006). Is a separate block of time for oral English language development in programs for English learners needed? . *Elementary School Journal, 107*(2), 181-198.

Scanlon, D. M., & Vellutino, F. R. (1996). Prerequisite skills, early instruction, and success in first-grade reading: Selected results from a longitudinal study. *Mental Retardation and Developmental Disabilities Research Reviews, 2*(1), 54-63.

Tabors, P. O., & Snow, C. E. (2001). Young bilingual children and early literacy development. In D. K. Dickinson & S. B. Neuman (Eds.), *Handbook of early literacy research* (Vol. 1, pp. 159-178). New York: Guilford Press.

Appendix B. Tables and Figures
Not included in page count.

Table 1: Counts of observations (n = 315 teachers, 942 total observations)

Program	Kindergarten			First Grade			Second Grade		
	Fall	Winter	Spring	Fall	Winter	Spring	Fall	Winter	Spring
Dual	16	16	16	26	26	26	20	21	22
Immersion	34	34	35	49	50	50	46	45	46
Transitional	34	35	35	44	45	46	42	41	42
Total	84	85	86	119	121	122	108	107	110

Table 2: Descriptive statistics of observed minutes in each language and instructional domain

		English Reading/Language Arts						Spanish Reading/Language Arts					
		Fall		Winter		Spring		Fall		Winter		Spring	
Grade	Program	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
K	Dual	6.3	8.1	19.5	20.9	22.6	30.9	49.4	28.0	57.2	36.7	39.7	15.1
	Immersion	37.2	23.5	36.7	24.0	46.5	29.2	2.4	8.0	1.3	5.4	1.3	4.0
	Transition	13.0	19.1	11.2	15.9	11.1	15.4	50.6	31.4	54.2	34.0	47.3	32.3
1	Dual	17.1	16.3	18.0	18.2	17.1	17.7	30.4	28.3	33.3	30.5	33.9	33.1
	Immersion	62.0	24.8	68.6	33.4	60.6	29.8	3.8	10.5	3.1	8.2	1.5	3.7
	Transition	17.9	23.7	18.8	20.5	22.0	24.8	58.7	30.2	57.5	33.0	50.3	30.1
2	Dual	42.2	35.5	36.3	36.8	32.1	28.7	40.8	27.2	42.5	31.4	34.0	29.1
	Immersion	79.2	31.5	77.5	28.1	69.5	32.9	3.6	15.3	5.0	21.4	6.0	19.7
	Transition	51.5	45.1	55.0	42.7	50.2	37.8	35.0	34.1	26.3	26.3	31.7	36.9

		English Oral Language Instruction						Spanish Oral Language Instruction					
		Fall		Winter		Spring		Fall		Winter		Spring	
K	Dual	11.8	13.9	15.3	16.2	9.3	11.8	16.4	11.4	15.0	13.9	8.8	11.3
	Immersion	20.1	15.9	18.4	13.4	15.6	12.1	0.5	1.5	0.3	0.9	0.7	2.1
	Transition	7.2	7.9	7.4	6.9	8.0	7.9	13.7	11.6	11.3	11.7	9.2	7.8
1	Dual	10.3	10.0	5.3	7.2	7.6	8.3	3.0	4.0	2.1	3.0	3.1	4.8
	Immersion	9.4	7.1	9.4	7.7	6.4	6.1	0.2	1.4	0.3	1.0	0.1	0.5
	Transition	6.5	7.4	4.9	9.1	5.6	7.3	5.8	6.6	4.5	6.5	4.9	6.4
2	Dual	1.9	3.7	2.6	3.7	2.0	5.1	1.4	2.8	1.3	2.5	2.2	5.6
	Immersion	5.7	7.8	4.6	8.4	4.0	5.1	0.5	2.6	0.4	1.5	0.5	2.4
	Transition	5.3	7.0	4.4	5.4	4.3	7.0	1.7	3.2	1.3	3.4	1.3	5.1

Table 3: English Reading and Language Arts Instruction, Mixed Effects Model Results for Grades K-2

Fixed Effect	Kindergarten				First Grade				Second Grade			
	b	SE	p		b	SE	p		b	SE	p	
Intercept	11.7	4.5	0.02		13.8	5.0	0.02		51.2	8.0	<.0001	
time	-0.5	1.6	0.78		2.2	1.7	0.18		-0.9	2.1	0.67	
Dual	-10.0	7.5	0.19		1.1	8.2	0.89		-2.8	13.5	0.83	
Immersion	22.9	5.8	<.01		51.4	6.7	<.01		33.1	10.8	<.01	
Transition[a]	0.0	—			0.0	—			0.0	—		
time*Dual	8.7	2.9	<.01		-1.8	2.7	0.51		-5.5	3.7	0.14	
time*Immersion	5.2	2.3	0.03		-2.9	2.3	0.21		-3.1	3.0	0.31	
time* Transition[a]	0.0	—			0.0	—			0.0	—		

Random Effect	Est.	SD	p	ICC	Est.	SD	p	ICC	Est.	SD	p	ICC
School Intercept	166.9	12.9	<.01	45%	117.1	10.8	0.02	22%	231.1	15.2	0.04	14%
Teacher Intercept	26.6	5.2	0.37	7%	164.2	12.8	0.08	31%	1,096.5	33.1	<.01	65%
Teacher covariance	2.9	—	0.93		17.2	—	0.70		-126.4	—	0.14	
Teacher Slope	23.5	4.8	0.10	6%	10.8	3.3	0.31	2%	37.9	6.2	0.13	2%
rater	20.2	4.5	0.18	5%	22.2	4.7	0.10	4%	37.4	6.1	0.11	2%
Residual	130.2	11.4	<.01	35%	215.6	14.7	<.01	41%	293.2	17.1	<.01	17%

Note: ICC = intraclass correlation. [a] The transitional program is the reference category for comparison.

Table 4: Spanish Reading and Language Arts Instruction, Mixed Effects Model Results for Grades K-2

Fixed Effect	Kindergarten				First Grade				Second Grade		
	b	SE	p		b	SE	p		b	SE	p
Intercept	53.3	5.7	<.01		65.1	4.7	<.01		35.9	5.3	<.01
time	-2.2	1.8	0.21		-4.0	1.5	<.01		-1.9	1.7	0.28
Dual	4.7	9.6	0.63		-34.1	7.9	<.01		6.7	9.2	0.47
Immersion	-47.8	7.6	<.01		-61.3	6.5	<.01		-33.0	7.3	<.01
Transition[a]	0.0	—			0.0	—			0.0	—	
time*Dual	-2.4	3.2	0.45		5.6	2.4	0.02		-0.3	3.0	0.92
time*Immersion	1.1	2.5	0.67		2.8	2.0	0.16		2.7	2.4	0.27
time* Transition[a]	0.0	—			0.0	—			0.0	—	

Random Effect	Est.	SD	p	ICC	Est.	SD	p	ICC	Est.	SD	p	ICC
School Intercept	149.9	12.2	0.04	25%	73.3	8.6	0.08	12%	33.8	5.8	0.27	5%
Teacher Intercept	208.0	14.4	<.01	34%	342.3	18.5	<.01	57%	462.9	21.5	<.01	62%
Teacher covariance	—[b]				—[b]				—[b]			
Teacher Slope	—[b]				—[b]				—[b]			
rater	39.0	6.2	0.11	6%	—[b]				—[b]			
Residual	207.2	14.4	<.01	34%	185.5	13.6	<.01	31%	250.8	15.8	<.01	34%

Note: ICC = intraclass correlation. [a] The transitional program is the reference category for comparison. [b] Random variability could not be estimated, so this parameter is set to zero.

Table 5: English Oral Language Instruction, Mixed Effects Model Results for Grades K-2

Fixed Effect	Kindergarten			First Grade			Second Grade		
	b	SE	p	b	SE	p	b	SE	p
Intercept	8.2	3.0	0.02	6.3	1.7	<.01	5.3	1.5	<.01
time	0.4	1.0	0.70	-0.5	0.6	0.39	-0.5	0.6	0.45
Dual	1.3	4.9	0.79	6.1	2.9	0.04	-3.0	2.6	0.26
Immersion	16.1	3.7	<.01	3.9	2.3	0.10	1.3	2.1	0.53
Transition[a]	0.0	—		0.0	—		0.0	—	
time*Dual	-1.6	1.9	0.39	-0.9	1.0	0.39	0.5	1.1	0.64
time*Immersion	-2.6	1.5	0.09	-1.0	0.8	0.24	-0.4	0.9	0.64
time* Transition[a]	0.0	—		0.0	—		0.0	—	

Random Effect	Est.	SD	p	ICC	Est.	SD	p	ICC	Est.	SD	p	ICC
School Intercept	75.3	8.7	<.01	49%	14.8	3.8	<.01	24%	4.9	2.2	0.03	12%
Teacher Intercept	—[b]				15.3	3.9	<.01	25%	5.0	2.2	0.03	12%
Teacher covariance	—[b]				—[b]				—[b]			
Teacher Slope	—[b]				—[b]				—[b]			
rater	5.4	2.3	0.23	4%	0.5	0.7	0.36	1%	—[b]			
Residual	73.2	8.6	<.01	48%	31.2	5.6	<.01	51%	31.9	5.6	<.01	76%

Note: ICC = intraclass correlation. [a] The transitional program is the reference category for comparison. [b] Random variability could not be estimated, so this parameter is set to zero.

Table 6: Spanish Oral Language Instruction, Mixed Effects Model Results for Grades K-2

Fixed Effect	Kindergarten				First Grade				Second Grade			
	b	SE	p		b	SE	p		b	SE	p	
Intercept	15.6	2.0	<.01		6.9	1.7	<.01		1.8	0.7	0.01	
time	-2.2	0.7	<.01		-0.5	0.3	0.11		-0.2	0.4	0.62	
Dual	5.4	3.5	0.12		-3.1	1.5	0.04		-1.0	1.2	0.43	
Immersion	-15.1	2.8	<.01		-4.8	1.3	<.01		-1.3	1.0	0.17	
Transition[a]	0.0	—			0.0	—			0.0	—		
time*Dual	-1.6	1.3	0.21		0.8	0.5	0.17		0.6	0.6	0.37	
time*Immersion	2.2	1.0	0.03		0.6	0.5	0.22		0.2	0.5	0.70	
time* Transition[a]	0.0	—			0.0	—			0.0	—		
Random Effect	**Est.**	**SD**	**p**	**ICC**	**Est.**	**SD**	**p**	**ICC**	**Est.**	**SD**	**p**	**ICC**
School Intercept	8.8	3.0	0.15	12%	2.4	1.5	0.05	4%	—[b]			
Teacher Intercept	29.8	5.5	<.01	40%	5.0	2.2	<.01	9%	8.1	2.9	<.01	50%
Teacher covariance	—[b]				—[b]				-4.1	—	<.01	
Teacher Slope	—[b]				—[b]				3.0	1.7	<.01	19%
rater	1.1	1.0	0.32	1%	38.7	6.2	0.02	70%	—[b]			
Residual	34.8	5.9	<.01	47%	9.3	3.1	<.01	17%	5.0	2.2	<.01	31%

Note: ICC = intraclass correlation. [a] The transitional program is the reference category for comparison. [b] Random variability could not be estimated, so this parameter is set to zero.