Routledge
Taylor & Francis Group

Journal of Research on Educational Effectiveness

ISSN: 1934-5747 (Print) 1934-5739 (Online) Journal homepage: https://www.tandfonline.com/loi/uree20

# Factors Contributing to Teachers' Sustained Use of Kindergarten Peer-Assisted Learning Strategies

**Devin M. Kearns , Douglas Fuchs , Kristen L. McMaster , Laura Sáenz , Lynn S. Fuchs , Loulee Yen , Coby Meyers , Marc Stein , Donald Compton , Mark Berends & Thomas M. Smith**

**To cite this article:** Devin M. Kearns , Douglas Fuchs , Kristen L. McMaster , Laura Sáenz , Lynn S. Fuchs , Loulee Yen , Coby Meyers , Marc Stein , Donald Compton , Mark Berends & Thomas M. Smith (2010) Factors Contributing to Teachers' Sustained Use of Kindergarten Peer-Assisted Learning Strategies, Journal of Research on Educational Effectiveness, 3:4, 315-342, DOI: 10.1080/19345747.2010.491151

**To link to this article:** https://doi.org/10.1080/19345747.2010.491151

*Journal of Research on Educational Effectiveness*, 3: 315–342, 2010

ISSN: 1934-5747 print / 1934-5739 online
DOI: 10.1080/19345747.2010.491151

Routledge
Taylor & Francis Group

**INTERVENTION, EVALUATION, AND POLICY STUDIES**

# Factors Contributing to Teachers' Sustained Use of Kindergarten Peer-Assisted Learning Strategies

**Devin M. Kearns and Douglas Fuchs**
Vanderbilt University, Nashville, Tennessee, USA

**Kristen L. McMaster**
University of Minnesota, Minneapolis, Minnesota, USA

**Laura Sáenz**
University of Texas, Pan American, Edinburg, Texas, USA

**Lynn S. Fuchs, Loulee Yen, Coby Meyers, Marc Stein, Donald Compton, Mark Berends, and Thomas M. Smith**
Vanderbilt University, Nashville, Tennessee, USA

**Abstract:** Factors were explored that predicted whether teachers sustained the use of a validated reading intervention. Seventy-three teachers from 37 schools in 3 states were asked in interviews whether they continued to use Kindergarten Peer Assisted Learning Strategies (KPALS) 1 year after their involvement in the program. A logistic regression model was created with teachers' yes/no responses as the dependent variable and with predictors identified as important to sustainability. Findings were consonant with current theoretical models of sustainability. The logistic regression model captured many key elements of teachers' decisions to sustain. Strongest predictors were teacher perceptions of the effectiveness of KPALS and degree of external technical support given them.

This paper was supported by funding from the U.S. Department of Education's Institute of Education Sciences, under grant number R305G040104. Funding for Devin Kearns was provided by an IES grant to Vanderbilt's ExpERT program for doctoral training (R305B040110).

All opinions expressed in this paper represent those of the authors and not necessarily the institutions with which they are affiliated or the U.S. Department of Education. All errors in this paper are solely the responsibility of the authors.

Address correspondence to Douglas Fuchs, Peabody #228, 230 Appleton Place, Vanderbilt University, Nashville, TN 37203, USA. E-mail: doug.fuchs@vanderbilt.edu

**Keywords:** Sustainability, evidence-based practice, education policy, scientifically based research

The use of rigorous research designs to evaluate educational programs has waxed and waned over several decades (Gersten, Baker, Smith-Johnson, Flojo, & Hagen-Burke, 2004). Such designs, particularly randomized control trials and quasi-experiments, are currently popular. They are important because they allow researchers to make causal claims regarding the effectiveness of their programs, something not possible with many other methods (Campbell & Stanley, 1963; Gersten et al., 2005; Shadish, Cook, & Campbell, 2003; Whitehurst, 2003).

Since the mid-1990s, many researchers have explored the effectiveness of peer-mediated instruction by using relatively rigorous designs—random assignment, measurement of the fidelity with which treatments are implemented, and appropriate statistical methods (e.g., Fuchs, Fuchs, Al Otaiba, et al., 2001; Fuchs, Fuchs, Mathes, & Simmons, 1997; Fuchs, Fuchs, Thompson, Al Otaiba, et al., 2001; Fuchs, Fuchs, Thompson, Svenson, et al., 2001; Greenwood, Delquadri, & Hall, 1989; Slavin et al., 1996; Vadasy, Jenkins, Antil, Phillips, & Pool, 1997). Fuchs and colleagues' Kindergarten Peer-Assisted Learning Strategies (KPALS), a supplemental reading program, is one such peer-mediated program. Evaluations of KPALS have been conservative as well as rigorous because the program is meant to be used only 30 min per day and three or four times a week; studies of it typically last only between 15 and 20 weeks. Moreover, teachers implement it after minimal professional development (usually 6 hr).

KPALS is lesson driven. Each lesson includes teacher-directed "Sound Play" phonological awareness activities plus teacher-led and peer-mediated practice with sound-symbol correspondence ("What Sound?"), word decoding and sentence reading ("Sound Boxes"), and sight word reading ("Sight Words"). See Figure 1 for a sample lesson and McMaster, Fuchs, and Fuchs (2006) for further description of the program. Despite its lesson-driven nature, KPALS is largely peer mediated, and its strength thus depends on pairs of novice readers (i.e., 5-year-olds) supporting each other effectively. Finally, all KPALS research has been conducted in schools, and instruction is delivered by teachers, considerations that make research findings more applicable to teachers but that also make the research itself more difficult to conduct.

Despite its abbreviated and supplementary nature, and that it has been implemented by teachers rather than by researchers in validation studies, students of teachers using KPALS have demonstrated significantly greater reading improvement than their counterparts in business-as-usual classrooms (Fuchs, Fuchs, Thompson, Al Otaiba, et al., 2001; Fuchs, Fuchs, Thompson, Svenson, et al., 2001; Fuchs & Fuchs, 2005). The KPALS studies, therefore, offer evidence that randomized control studies in real-world settings can produce

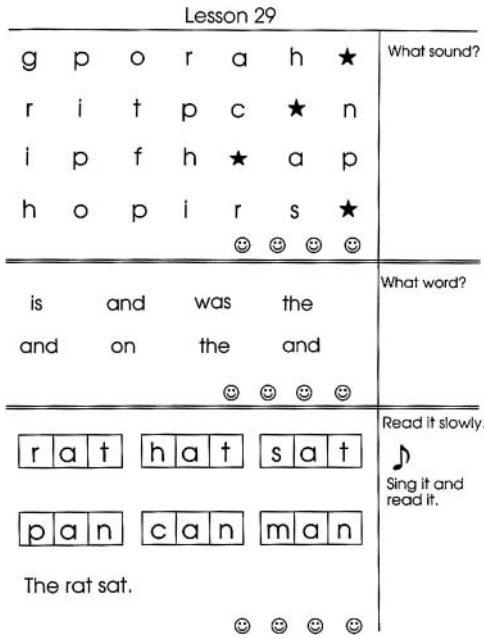


***Figure 1.*** A sample KPALS lesson. *Note*. The top section shows "What Sound," designed to practice sound-symbol correspondence. The middle section shows "Sight Words," the sight word activity. The bottom section shows "Sound Boxes," which includes decoding practice with Elkonin boxes and sentence reading. All three are practiced as teacher-directed and peer-mediated activities. Smiley faces are part of a reinforcement system. They are marked each time one student in the pair completes the activity. Students also read short books in pairs after completing the main lesson.

educationally significant effects. The application of rigorous methods also gives us greater confidence that obtained effects are meaningful.

## THE RESEARCH–PRACTICE GAP

The KPALS studies and others' research (e.g., Borman et al., 2006; O'Connor, 1999; Slavin, Madden, & Datnow, 2007) suggest that rigorous methods can produce positive effects in which we have confidence. But the development and empirical validation of such practices do not mean that everyone—or anyone—uses them. Indeed, evidence suggests the existence of a persistent research–practice gap, that research is not successfully reaching classrooms (e.g., Boardman, Arguelles, Vaughn, Hughes, & Klingner, 2005; Bos, Mather, Dickson, Podhajski, & Chard, 2001; Elmore, 1996; Gersten, Vaughn, Deshler, & Schiller, 1997).

One major cause of this gap has been the historical lack of demand for research-based educational change (Carnine, 1997; Fuchs & Fuchs, 1998) and a concomitant absence of a competitive market for evidence-backed instructional procedures, curricula, and materials. Although teachers and administrators certainly want their students to achieve acadmically, their use of empirically validated practices has been inconsistent (Greenwood & Abbott, 2001; Klingner, Ahwee, Piloneta, & Menendez, 2003; McLaughlin & Mitra, 2001; Tyack & Tobin, 1994). One important precondition for closing the research–practice gap, therefore, must be a demand for the adoption of evidence-based practices. Given the current emphasis on content standards and ambitious achievement goals, the once-lamented absence of supply (Berends, Chun, Schuyler, Stockly, & Briggs, 2002) and demand (Carnine, 1997) is no more.

A climate where evidence-based instructional programs are promoted and available does not, however, mean teachers and administrators are eager to adopt them. Evidence-based programs tend to be prescriptive because the need for experimental control dictates that treatments be highly specified. This, in turn, requires exact teacher and student behaviors. Ball (1995) and Darling-Hammond (1996) have argued that the directive nature of many research-validated programs may cause many teachers to view them as inflexible and disrespectful of their craft knowledge. If Ball and Darling-Hammond are right, then the growing number of research-backed programs in a climate promoting their use may be immaterial: They will sit unused on classroom shelves. Researchers who overlook the role of teacher and administrator attitudes toward evidence-based programs often find that their attempts at scaling up fall short: Teachers unconvinced of a program's value are likely to implement it half-heartedly, if at all.

The authors recently conducted a 5-year scaling-up study of PALS with the first 2 years focusing on KPALS (2004–2005 and 2005–2006). They chose three sites in which educators had different prior experiences with the program: Tennessee, a great deal; Minnesota, some; and Texas, virtually none. The investigators examined the research–practice gap using fidelity of treatment implementation as a multidimensional indicator of willingness and ability to implement KPALS (cf. Dusenbury, Brannigan, Falco, & Hansen, 2003). The research team designed three levels of increasingly strong teacher support to determine whether greater support would mean greater fidelity of KPALS implementation. The levels of support were (a) KPALS Workshop (KPALS-W), where teachers were given the KPALS manual and materials and participated in a 6-hr workshop; (b) KPALS-Workshop + Booster (KPALS-W+B), where teachers attended the workshop plus three 1-hr "booster sessions" in which they interacted with researchers and other teachers to refine their implementation; and (c) KPALS-Workshop + Booster + Helper (KPALS-W+B+H), where teachers participated in the workshop and booster sessions and received the support of a research assistant 2 days per week. Teachers were randomly assigned within schools to one of these conditions or control, in which case they continued instruction as usual.

Initial findings suggested that KPALS promoted phonological awareness across the three sites. It was particularly effective in Nashville, where KPALS students outperformed controls on most phonological awareness, alphabetics, and fluency measures (Fuchs et al., 2008). Also in Nashville, more intensive levels of teacher support promoted higher levels of teacher and student fidelity of treatment implementation, stronger letter-sound correspondence, and more fluent reading in connected text. These results suggest that greater—but not heroic—external technical support from research teams may increase teacher capacity to implement evidence-based practices, which in turn may improve teacher attitudes toward them. In other words, providing support that improves implementation may bridge the gap between research and practice.

## BEYOND THE RESEARCH–PRACTICE GAP

If we close the research–practice gap for a single year only to find teachers have moved on to something else in the next year, then we are only playing Sisyphus. There is a long history of teachers reverting to prior practices (Datnow, 2002; Giles, 2006; McLaughlin & Mitra, 2001; Tyack & Tobin, 1994). Therefore, it is necessary to identify what makes teachers *sustain* effective practices.

The experimental design of the most recent KPALS study provided an opportunity to examine the impact of external technical support on sustainability. After the 2005–2006 school year, support from our research team ended for the KPALS teachers, but we contacted them again about a year later, in February and March of 2007. Research assistants spoke with 66% of teachers and asked them the following question: "In the 2006–2007 school year, have you continued to use KPALS?" Combined with information about their study condition, we were able to explore how external technical support and other factors may have influenced their decisions to sustain the use of KPALS. Before sharing these results, we present a sustainability model we constructed from the extant literature and subsequently used to make sense of our sustainability data.

## SUSTAINABILITY

Sustainability is simply the likelihood a program will be used for the long term. The dimensions of its nature (the factors that increase or decrease the likelihood of sustaining), however, are more complex. After providing our operational definition of sustainability, we describe several dimensions theorized to underlie it. We also describe how we measured each of these constructs in the present study.

### Definition

Datnow (2002) described sustainability as longevity or institutionalization, the latter being the integration of a practice into "business as usual." But teachers

often add to, subtract from, or otherwise modify programs (Coburn, 2003; Datnow, Hubbard, & Mehan, 1998; McLaughlin & Mitra, 2001), creating adaptations quite different from their validated antecedents. Han and Weiss (2005) opted, therefore, for a more ambitious sustainability standard, arguing that practices are sustained only when they are continued with "fidelity to the core program elements" (p. 666). We operationalize sustainability, therefore, as the continued implementation of a program with fidelity to the core elements after external supports are no longer present.

## Dimensions of Sustainability

Decisions to sustain occur at multiple levels (teacher, building, local education agency, state, federal; Datnow, 2005), but we focus on the teacher's decision to sustain. The factors in teacher decision making are most relevant in this study because most KPALS teachers were not required to continue the program (i.e., most schools and districts did not mandate its ongoing use). We propose a model, shown in Figure 2, of factors related to teacher decisions to continue program implementation. The model is anchored by four "dimensions" described in the extant literature on sustainability: *external technical support, implementation experiences, teacher characteristics*, and *school/district support*. Each dimension is understood to comprise underlying constructs, which we call "subdimensions." We describe this sustainability-prediction model, but we wish to underscore its heuristic intent.

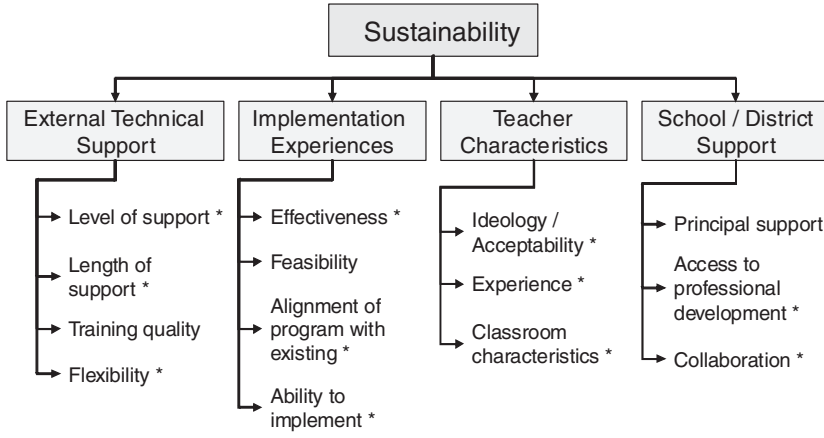


***Figure 2.*** Theory-identified factors predicting a teacher's independent decision to sustain the use of an educational reform. *Note*. Items with asterisks included in the present analysis.

*External Technical Support.* We define external technical support as assistance to teachers provided by external developers to assure a program is implemented successfully. Although early research suggested a negative relationship between external technical support and sustainability (Rand Corporation, 1978), subsequent research suggests external technical support improves the likelihood of sustainability (Desimone, 2003; Gersten, Chard, & Baker, 2000; Glennan, Bodilly, Galegher, & Kerr, 2004; Klingner et al., 2003; McLaughlin, 1990). Researchers have suggested the following subdimensions of external technical support: *a high level of support*, *lengthier support*, and *high-quality training*.

In addition, some have argued that the *flexibility* developers permit teachers in implementation predicts sustainability. As we previously suggested, the emphasis on fidelity in practices backed by scientifically-based research may be distasteful to teachers (Ball, 1995), and they may opt to drop the practice once external developers leave. Han and Weiss (2005) suggested, to the contrary, that programs backed by scientifically based research build teachers' confidence because they are known to work, and they argued external technical support improves sustainability when focused on fidelity. Rose and Church (1998) reviewed studies of teacher training and found that fidelity-focused feedback produced short-term improvement in implementation, but they could not detect a conclusive relation with sustainability. The literature shows both ambivalence and ambiguity. We are especially interested, therefore, in examining the role of flexibility in sustainability of KPALS.

*Implementation Experiences.* The second dimension of sustainability is teachers' implementation experiences. The *perceived effectiveness* subdimension means that teachers sustain programs they consider successful. Gersten et al. (2000) amended this idea by suggesting it is the *effectiveness of the program for low-achieving students* that drives teacher decisions to sustain. Whether a program seems *feasible* is also important (Gersten & Dimino, 2001; Han & Weiss, 2005; Vaughn, Hughes, Schumm, & Klingner, 1998). Datnow et al. (1998) observed that once program developers no longer propped up less feasible programs with resources and encouragement, implementation quickly dissipated. *Alignment with existing programs* is also relevant to sustainability because a gap between reform efforts and schools' present practices is likely to inhibit teachers' individual sustainability efforts (Datnow & Stringfield, 2000; Desimone, 2003). Finally, a teacher's willingness to sustain a program may be shaped by an *ability to implement* it, which differs from feasibility. Even if teachers think a program is feasible and has positive benefits for students, they may still discontinue its use if they feel unable to implement it well. Conversely, teachers who were successful implementers might be "inoculated" against giving up in the face of challenges in subsequent years (Han & Weiss, 2005, p. 676).

*Teacher Characteristics.* The third dimension, teacher characteristics, differs from the second because it relates to features of a teacher's experience apart

from KPALS but still relevant for the decision to sustain. *Personal ideology*, for example, may play a pivotal role in sustainability decisions (Datnow & Castellano, 2000; Datnow & Stringfield, 2000; Rohrbach, Graham, & Hansen, 1993; Rose & Church, 1988). The acceptability of a program, vis-à-vis a teacher's ideology, may play a key role in sustainability decision making. The literature also suggests the possibility that *teaching experience* negatively correlates with sustainability, on grounds that experienced teachers may be less flexible and more likely to revert to prior practices when programs end (Giles, 2006; Hargreaves & Goodson, 2006; Tyack & Tobin, 1994). Finally, *classroom characteristics*, particularly demographic characteristics of students may play an important role in teachers' decisions to sustain. For example, teachers may base decisions about using programs on whether students are English Language Learners (Stringfield, Datnow, & Ross, 1998). Prior research with KPALS suggests particularly strong effects for English Language Learners (McMaster, Kung, Han, & Cao, 2007; Sáenz, Fuchs, & Fuchs, 2005), so it is possible that teachers of these students sustained KPALS at higher rates.

*School and District Support.* Finally, school and district support may affect a teacher's decision to sustain. The subdimensions extensively discussed in the literature (e.g., Datnow & Stringfield, 2000; Desimone, 2003; Elmore, 1996; Glennan et al., 2004) are *principal support, access to professional development*, and *collaboration with other teachers*. The impact of principal support for the reform appears important because teachers' decisions are often shaped by them, even when implementation is not required. Access to professional development reflects the culture of the school and the district; districts that provide more professional development experiences for teachers are more likely to have teachers sustain innovative programs. Finally, the collaborative culture of the school may influence sustainability because a supportive climate may inspire teachers to continue a program, even when they face other demands on their time.

The four dimensions of sustainability in our model (external technical support, implementation experiences, teacher characteristics, and school and district support) reflect current thinking on sustainability in the literature. Not all models include all of the just-mentioned dimensions. But our relatively inclusive and heuristic model reflects the breadth of perspectives on factors influencing sustainability. These dimensions are the focus of both our research questions and analyses.

## Research Questions

Given the importance of understanding factors that influence sustainability and, more specifically, the role of external technical support in our study, we ask the following questions:

1. How well does our sustainability model predict overall whether teachers sustained KPALS? Does a reduced theoretical model capture sustainability as well as the complete model?
2. Which model components have the greatest influence on teachers' decisions to sustain KPALS?
3. Does the level of external technical support provided by our research team influence the likelihood of sustainability?
4. Which is the better predictor of sustainability: teachers' perceived effectiveness of KPALS for all students or the perceived effectiveness of KPALS for low-achieving students?

## METHOD

### Sample

Seventy-three kindergarten teachers from 40 schools in Tennessee, Minnesota, and Texas agreed to participate in interviews to discuss their implementation with us. The teachers were a subset of the original 134 teachers who participated in the KPALS study during 2005–2006 (see Table 1). From the original sample, control teachers ($n = 24$) were not eligible for interviews. Of the remaining 110 KPALS teachers, we did not interview teachers who could not teach KPALS because they changed grades ($n = 9$) or were no longer teaching ($n = 6$). This reduced the eligible number of teacher respondents to 95. Sixteen teachers did not respond to our invitation, and 5 opted not to participate. One of the teachers we interviewed was omitted from analysis due to extensive missing data. Therefore, 73 of 95 teachers, or 77% of the eligible respondents, were included in analysis.

Chi-square analyses and one-way analyses of variance, shown in Table 2, revealed no statistically significant differences between the interviewed ($n = 73$) and noninterviewed ($n = 37$) KPALS teachers on teacher demographic

**Table 1.** Teachers included and not included in the study

| | Inclusion Status | | | | | |
|---|---|---|---|---|---|---|
| Treatment | Included | Not Eligible | Declined | No Response | Missing Data | Total |
| KPALS+W | 32 | 7 | 2 | 6 | 0 | 47 |
| KPALS+W+B | 30 | 5 | 2 | 9 | 1 | 47 |
| KPALS+W+B+H | 11 | 3 | 1 | 1 | 0 | 16 |
| Total | 73 | 15 | 5 | 16 | 1 | 110 |

*Note*. KPALS = Kindergarten Peer-Assisted Learning Strategies; W = Workshop; B = Booster; H = Helper.

**Table 2.** Descriptive and group-equivalence statistics for teacher demographics and survey responses

| | PALS Teacher and Teacher Classroom Data by Interview Status | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Interviewed | | | | Not Interviewed | | | | | |
| Variable | *M* | *SD* | *N* | % | *M* | *SD* | *N* | % | $F^a$ | $\chi^2$ |
| *Demographic variables not in analysis* | | | | | | | | | | |
| Gender (female) | | | 71 | 97.3 | | | 37 | 97.3 | | 0.01 |
| Race | | | | | | | | | | 0.70 |
| Black | | | 8 | 11.0 | | | 4 | 10.8 | | |
| White | | | 42 | 57.5 | | | 24 | 64.9 | | |
| Hispanic | | | 21 | 28.8 | | | 8 | 21.6 | | |
| Other | | | 2 | 2.7 | | | 1 | 2.7 | | |
| *Variables used in analysis* | | | | | | | | | | |
| Sustained KPALS | | | | | | | | | | |
| Site | | | | | | | | | | 0.12 |
| Tennessee | | | 30 | 41.1 | | | 16 | 43.2 | | |
| Minnesota | | | 21 | 28.8 | | | 11 | 29.7 | | |
| Texas | | | 22 | 30.1 | | | 10 | 27.0 | | |
| Treatment condition | | | | | | | | | | 0.24 |
| Workshop | | | 32 | 43.8 | | | 15 | 40.5 | | |
| Booster | | | 30 | 41.1 | | | 17 | 46.0 | | |
| Helper | | | 11 | 15.1 | | | 5 | 13.5 | | |
| Sustained KPALS | | | 52 | 71.2 | | | | *NA* | | |
| Returning teacher | | | 42 | 57.5 | | | 12 | 32.4 | | 6.19* |
| Perceived eff. | 4.57 | 0.39 | | | 4.38 | 0.43 | | | 5.40* | |
| Alignment | 2.97 | 0.80 | | | 3.09 | 0.56 | | | 0.63 | |
| % fidelity | 0.86 | 0.10 | | | 0.81 | 0.14 | | | 3.82† | |
| Phonics per. | | | 37 | 50.7 | | | 20 | 54.1 | | 0.11 |
| Whole language per. | | | 15 | 20.6 | | | 6 | 16.2 | | 0.30 |
| Experience | 12.96 | 8.85 | | | 11.76 | 8.31 | | | 0.47 | |
| ELL students[b,c] | 0.25 | 0.34 | | | 0.15 | 0.29 | | | 2.56 | |
| Principal support | 2.61 | 0.58 | | | 2.59 | 0.59 | | | 0.03 | |
| Literacy PD factor | 0.00 | 0.89 | | | 0.12 | 0.78 | | | 0.46 | |
| Collaboration factor | 0.00 | 0.93 | | | −0.01 | 1.02 | | | 0.00 | |
| Perceived eff. (low) | 4.70 | 0.50 | | | 4.44 | 0.65 | | | 5.17* | |

*Note*. KPALS = Kindergarten Peer-Assisted Learning Strategies; perceived eff. = perceived effectiveness; per. = perspective; Alignment = teacher perception of alignment between KPALS and school programs; Experience = teaching experience, in years; ELL = English language learner; PD = professional development; (low) = low students.

[a]$df = (1, 108)$. [b]In year of study participation. [c]Measured as a proportion.

†$p < .10$. *$p < .05$.

variables. Among the predictors of sustainability, however, interviewed and noninterviewed teachers differed significantly in terms of (a) whether they were returning teachers, $\chi^2(1, N = 110) = 6.191, p < .01$; (b) their perception of KPALS effectiveness, $F(1, 108) = 5.40, p = .02$; and (c) their perception of KPALS effectiveness for low-achieving students, $F(1, 108) = 5.17, p = .03$. Interviewed teachers were more likely to be returning teachers and to view KPALS as effective for all students and low-achieving students.

## Data Collection and Variables Used in Analysis

The data presented in this article were derived from three sources. The first was the aforementioned teacher interviews conducted in 2007, the year after the KPALS study ended. The face-to-face interviews were undertaken by 25 graduate students employed as research assistants across the three sites, who trained with one of the authors for 2 hr prior to interviewing the teachers. Most of the interviews ($n = 65$) lasted 30 to 45 min and included questions about the KPALS activities teachers used and how they used them as well as open-ended questions about their reasons for continuing or discontinuing use of KPALS. Eight teachers were not available for in-person interviews but participated in phone interviews in which they were asked only whether they were continuing to use KPALS. The second data source was a survey that teachers completed during the 2005–2006 school year, when they implemented KPALS. The survey was developed by the research team, using items from prior national surveys, including the National Assessment of Educational Progress, National Longitudinal Study of No Child Left Behind, and Schools and Staffing Survey and several researcher-developed items. The third data source was a teacher demographic form completed by the teachers during the 2005–2006 year.

The sustainability outcome was defined by teacher responses to the interview question, "Have you continued to use KPALS?" Responses were coded dichotomously: 1, if they reported continued use of KPALS; 0, if they said they did not continue its use. To check the accuracy of these yes–no self-reports, we used the data from the 65 teachers who participated in the longer in-person interview. We asked questions about which KPALS activities they used, how they implemented them, and with what frequency and duration. Two of the authors coded these teacher responses and entered them in a latent variable model that used the coded data to classify teachers as sustaining or not. For 63 of the 65 teachers, the latent variable prediction proved to be the same as the yes–no response. This strong alignment ($r = .99$) between the latent variable prediction and the yes–no self-report in the interview indicates the yes–no responses are a reasonable reflection of teacher sustainability. Thus, the latent variable model justified the use of the full sample of 73 teachers that included teacher responses from the shorter phone interviews. For the present study,

therefore, the dependent variable was the yes–no report, not any product of the latent variable analysis.

We then chose predictors that represented the subdimensions of our four-dimension sustainability model (see Figure 2). Fifteen variables were created, reflecting all the subdimensions except training quality and program feasibility, for which we had no data. We describe these predictors next.

*External Technical Support.* We operationalized the four subdimensions of external technical support in the following way. Level of support was examined in terms of our treatment conditions, *KPALS-W*, *KPALS-W+B*, and *KPALS-W+B+H* (the latter only present at the Tennessee site). The last two levels were used as dichotomous variables in the analysis; KPALS-W was a reference category. For the length of support, *returning* teachers—that is, those who were KPALS teachers in 2004–2005 and 2005–2006—had received support over a longer period. Of 73 interviewed teachers, 42 were returning teachers; we coded this variable dichotomously (1 = returning). We did not collect data regarding teacher satisfaction with the Workshop, Booster sessions, or Helper, so we could not model training quality. Flexibility was captured in a limited way through the treatment conditions, because teachers in the Workshop were permitted de facto flexibility by virtue of having no additional support. In addition, *average treatment fidelity* measured by the researchers at two different times reflected flexibility in that teachers with lower fidelity sometimes made adaptations that reduced their fidelity ratings.

*Implementation Experiences.* We operationalized *perceived effectiveness* using our teachers' survey responses. Teachers rated the effectiveness of KPALS for high-, average-, and low-achieving students and all students. The scale comprises 21 items and has an alpha reliability of .86. *Perceived effectiveness for low-achieving students* was measured using only the items from the same survey related to low-achieving students ($\alpha = .85$). *Teacher perception of alignment of KPALS* with existing programs was a second variable created using teacher responses to a single item that asks teachers to rate whether they "strongly agree," "agree," "disagree," or "strongly disagree" with the statement, "Reading PALS is an example of a program that exhibits continuity with other programs at this school." The 4-point scale ranged from 1 (*strongly disagree*) to 4 (*strongly agree*). Feasibility was not modeled because we did not have the requisite data.

*Teacher Characteristics.* To explore the subdimension of teacher ideology, we asked teachers on the survey the following question: "Do you consider yourself more of a whole-language teacher or a phonics teacher?" Teachers could respond, "whole language," "phonics," or "both." The *whole language perspective* and the *phonics perspective* were coded as separate dichotomous variables, with "both" as the reference category. This was considered a good

measure of ideology and acceptability of the program because attitudes toward phonics and whole language often reflect teacher philosophy (Chall, 2000; Pressley, 2006). *Teacher experience* was the number of years the teacher had taught, including the study year. The only classroom characteristic was *percentage of students who were ELLs* during the 2005–2006 school year, which was the study year. Although other demographic characteristics were plausibly linked to sustainability (e.g., student achievement, behavior, risk for reading difficulty), we included only ELL status because this characteristic was known to be relevant in prior KPALS research (Saénz et al., 2005).

*School/District Support.* The variables supporting this dimension were derived from survey items regarding teacher perceptions of their school. *Teacher perception of principal support* was a scale derived from 17 survey items about the principal's leadership, such as, "To what extent do you agree or disagree with the following statement about your principal's leadership? The principal at my school talks with me frequently about Reading PALS." The alpha reliability of this scale was .95. The amount of *professional development* teachers received in literacy was created using principal components analysis for 37 professional development-related survey items. The factors were rotated using varimax rotation, resulting in seven factors with Eigenvalues greater than 1. The items related to literacy professional development created a factor with an Eigenvalue of 2.65, representing 13% of the total variance in the professional development variables. The *degree of collaboration* teachers perceived within their school was a variable created out of the same principal components analysis used to create the literacy professional development factor. This factor had an Eigenvalue of 3.20, representing 15% of the total variance in the professional development variables.

Table 2 includes descriptive statistics for included variables, and Table 3 shows their zero-order correlations. Many variables were derived from teacher surveys. Information about the surveys is available from the first author.

In addition to the variables in our theoretical model, we added variables for the three implementation sites (i.e., Tennessee, Minnesota, and Texas, with Tennessee as a reference category) to control for site characteristics not otherwise explained by variables in the model. We anticipated that "site" would be important because it was important in the fidelity models created by Stein et al. (2008) using the same data. Part of the site effect reflects varying experience with KPALS, with Tennessee having the most and Texas the least.

## Complete and Reduced Analytic Models

As already indicated, the outcome measure of whether teachers sustained KPALS was dichotomized such that *yes* was represented by 1 and *no* was represented by 0. We used logistic regression to estimate the variables' respective

**Table 3.** Correlation matrix for logistic regression predictors

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 Sustained KPALS | 1.00 | | | | | | | | | | | | | | | | |
| 2 PALS + Booster | .10 | 1.00 | | | | | | | | | | | | | | | |
| 3 PALS + Booster + Helper | −.32 | −.35 | 1.00 | | | | | | | | | | | | | | |
| 4 Returning | .25 | .04 | −.03 | 1.00 | | | | | | | | | | | | | |
| 5 Effectiveness | .07 | .00 | .20 | .07 | 1.00 | | | | | | | | | | | | |
| 6 Alignment | .07 | −.03 | .02 | .12 | −.05 | 1.00 | | | | | | | | | | | |
| 7 Fidelity | .13 | .15 | .24 | .18 | .17 | −.06 | 1.00 | | | | | | | | | | |
| 8 Phonics per. | −.14 | .21 | .03 | .04 | .01 | .03 | −.00 | 1.00 | | | | | | | | | |
| 9 Whole language per. | −.05 | −.22 | −.02 | .09 | .01 | −.02 | .08 | −.52 | 1.00 | | | | | | | | |
| 10 Experience | −.15 | .10 | .19 | −.06 | .20 | −.11 | .04 | .31 | −.23 | 1.00 | | | | | | | |
| 11 % ELL | .21 | −.04 | −.08 | .08 | −.23 | −.03 | −.01 | .07 | .03 | −.22 | 1.00 | | | | | | |
| 12 Principal support | .23 | .03 | .01 | .07 | −.08 | .12 | −.16 | .10 | −.12 | −.03 | .18 | 1.00 | | | | | |
| 13 PD access | .18 | .36 | −.09 | .03 | .08 | .08 | .05 | .08 | −.21 | −.19 | .01 | .00 | 1.00 | | | | |
| 14 Collaboration | .15 | −.08 | −.17 | −.11 | −.04 | −.01 | −.05 | .03 | −.04 | −.24 | −.00 | .23 | .07 | 1.00 | | | |
| 15 Minnesota | .27 | .02 | −.27 | .06 | −.22 | .10 | −.27 | −.22 | −.02 | .10 | −.09 | −.10 | .09 | −.14 | 1.00 | | |
| 16 Texas | −.11 | .06 | −.28 | −.10 | .21 | −.07 | −.04 | .11 | .11 | −.07 | .14 | .02 | −.06 | −.20 | −.42 | 1.00 | |
| 17 Effectiveness (low) | −.15 | −.00 | .12 | −.18 | .67 | −.13 | .03 | −.10 | .04 | .27 | −.40 | −.08 | −.05 | −.05 | −.09 | .29 | 1.00 |

*Note*. Correlations > .23 significant at $p < .05$. Correlations > .31 significant at $p < .01$. Correlations > .49 significant at $p < .001$. KPALS = Kindergarten Peer-Assisted Learning Strategies; Alignment = teacher perception of alignment between KPALS and school programs; per. = perspective; Experience = teaching experience, in years; ELL = English language learner; PD = professional development; (low) = low students.

contributions to a teacher's decision to continue KPALS or not. For binary outcomes in logistic regression, the probability of "success," $\varphi_{ij}$ —in this case, the probability KPALS will be sustained—is given by the link function

$$\varphi_{ij} = \frac{1}{1 + \exp\{-\eta_{ij}\}},$$

where $\eta_{ij}$ represents the log odds that KPALS will be sustained.

We began our data analysis with an unconditional hierarchical logistic regression model because teachers were nested within schools, with an average of about two per school (minimum = 1, maximum = 6). We tested whether there were significant school-level random effects using a likelihood ratio test. The effects were not significant (estimated variance = 1.35 [$SE = 0.75$]; $\chi$-bar$^2 = 2.32$; $p = .06$). We conducted the remaining regression analyses without school-level random effects.

We then conducted data analysis in three steps. First, we entered into the model all of the theoretically relevant variables for which we had data (indicated by stars in Figure 2). The resulting model is considered the *complete theoretical model*, although we could not model quality of support and feasibility because we did not collect these data. The complete theoretical model contained 15 predictors. Given the size of the sample and the number of teachers who did not sustain ($n = 23$), this model may have caused overfitting (Peduzzi, Concato, Kemper, Holford, & Feinstein, 1996). In the second step, we examined results from the complete theoretical model and removed variables with $p$ values above .25, following the recommendation of Hosmer and Lemeshow (1989). This cutoff was set high enough to prevent the elimination of those that might have a significant effect once the random noise associated with the nonsignificant variables was removed. We designated the new model the *reduced theoretical model* and ran the regression again. Our final step was to use the reduced theoretical model, replacing the perception of effectiveness for all students with the perception of effectiveness for low-achieving students. All results are reported in Table 4.

## RESULTS

### Complete Theoretical Model

The complete model allowed us to answer our first question: Using extant knowledge, can we predict sustainability? To answer this question, we examined the rate of correct classification, sensitivity (rate of correctly predicting sustaining teachers), and specificity (rate of correctly predicting nonsustaining teachers) using a probability cutoff of .50; that is, a teacher whose predicted probability of sustaining was .51 would be considered sustaining whereas

**Table 4.** Results from logistic regression models predicting the likelihood that kindergarten teachers sustained KPALS

| | Logistic Regression Models | | |
|---|---|---|---|
| | Complete Theoretical | Reduced Theoretical | Reduced + Effectiveness (Low) |
| Goodness of fit statistics | | | |
| Correctly classified[a] | .89 | .93 | .85 |
| Pearson $\chi^2$ | 33.89 | 43.35 | 75.95$^{\dagger}$ |
| Hosmer-Lemeshow $\chi^2$ | 0.40 | 0.51 | 7.72 |
| Sensitivity[a] | .94 | .96 | .90 |
| Specificity[a] | .76 | .86 | .71 |
| Regression coefficients | | | |
| Intercept | −115.45 (55.81)* | −73.93 (25.23)* | −15.95 (7.39)* |
| External support | | | |
| KPALS Workshop + Booster | −4.07 (1.92)$^{\dagger}$ | −2.63 (1.26) | −0.99 (0.97)$^{\dagger}$ |
| KPALS Workshop + Booster + Helper | −17.11 (7.42)* | −12.76 (4.61)** | −4.42 (1.81)* |
| Returning | 0.23 (1.30) | | |
| Implementation experiences | | | |
| Effectiveness | 11.97 (6.01)* | 7.76 (2.61)** | |
| Effectiveness (Low) | | | −0.02 (1.06) |
| Alignment | 1.01 (1.02) | | |
| Fidelity | 56.86 (26.19)* | 39.57 (14.88)** | 16.21 (5.81)** |
| Teacher characteristics | | | |
| Phonics per. | −3.49 (1.88)$^{\dagger}$ | −2.02 (1.45) | −1.29 (0.98) |
| Whole language per. | −7.44 (3.40)* | −6.15 (2.55)* | −1.84 (1.23) |
| Experience | 0.09 (0.10) | | |
| % ELL | 12.70 (6.88) | 5.64 (2.38)* | 1.90 (1.36) |
| School/District support | | | |
| Principal support | 5.67 (2.47)* | 4.30 (1.71)* | 1.93 (0.81)* |
| Literacy PD factor | −1.72 (1.12) | −0.06 (0.42) | −0.09 (0.26) |
| Collaboration | 2.00 (1.24) | 0.85 (0.59) | 0.38 (0.43) |
| Site | | | |
| Minnesota | 9.35 (5.36)$^{\dagger}$ | 6.26 (2.89)* | |
| Texas | −4.07 (1.92)* | −2.10 (1.26)$^{\dagger}$ | −2.43 (1.46)$^{\dagger}$ |

*Note*. Regression coefficients given with standard errors in parentheses. KPALS = Kindergarten Peer-Assisted Learning Strategies; (low) = low students; Alignment = teacher perception of alignment between KPALS and school programs; per. = perspective; Experience = teaching experience, in years; ELL = English language learner; PD = professional development;

[a]Values are probabilities.

$^{\dagger}p < .10$. $^{*}p < .05$. $^{**}p < .01$.

one with a predicted probability of .49 would be considered nonsustaining. The Pearson and Hosmer-Lemeshow chi-square statistics were not statistically significant ($p > .50$ for both tests), suggesting good model fit. Of the 73 teachers in our sample, 71% sustained. So we expected the complete model to have higher sensitivity than specificity, and this was the case: Sensitivity was .94, and specificity was .76. The complete model, then, effectively predicted the likelihood of sustainability. Figure 3 (top) shows the sensitivity and specificity of the model at different probability cutoffs.

Ten coefficients in the complete model were statistically significant at the .10 level, and of these, 6 were significant at the .05 level. This latter group was associated with the following variables: (a) KPALS-W+B+H, (b) perceived effectiveness of KPALS, (c) fidelity, (d) a whole language perspective, (e) principal leadership, and (f) being at the Texas site. Of the 15 predictors, 3 (participating in study for a 2nd year, alignment of KPALS with existing programs, and years of teaching experience) had $p$ values above .25 and were eliminated for the reduced theoretical model.

## Reduced Theoretical Model

In the reduced model the Hosmer-Lemeshow and Pearson chi-square statistics were both nonsignificant, although slightly larger than in the complete model. The fit was strong. Using the same .50 cutoff, the rate of correct classification was actually greater than with the complete model (.93 compared with .89; see Figure 4). Sensitivity and specificity were also somewhat higher (.96 vs. .94 and .86 vs. .76, respectively; see Figure 3, bottom). A comparison of the chi-square likelihood ratios for the complete model, $\chi^2(15, N = 73) = 57.11$, and the reduced model, $\chi^2(12, N = 73) = 54.35$, was not significant, $\chi^2$ difference $(3) = 2.76$, suggesting the removal of some variables did not significantly affect precision. Thus, the reduced theoretical model predicted sustainability as well as the complete model.

In addition, with the reduced model, we were able to detect statistically significant effects more easily and answer the remaining research questions with greater confidence. With regard to our second question (Which components of the model had the greatest influence on teachers' decisions to sustain?), we judged "influence" by the size of the $t$ statistics given for each of the coefficients. In order of influence, the significant predictors were (a) perception of KPALS effectiveness (7.76 [$SE = 2.61$], $t = 2.98$), (b) participation in the KPALS-W+B+H condition ($-12.76$ [$SE = 4.61$], $t = -2.77$), (c) average fidelity (39.57 [$SE = 14.88$], $t = 2.66$, (d) principal leadership (4.30 [$SE = 1.71$], $t = 2.52$), (e) expressing a whole language perspective ($-6.15$ [$SE = 2.55$], $t = -2.42$, (f) percentage of ELL students in the prior year (5.64 [$SE = 2.39$], $t = 2.36$), and (g) working at the Minnesota site (6.26 [SE = 2.89], $t = 2.16$). These results suggest that participating in the KPALS-W+B+H condition and

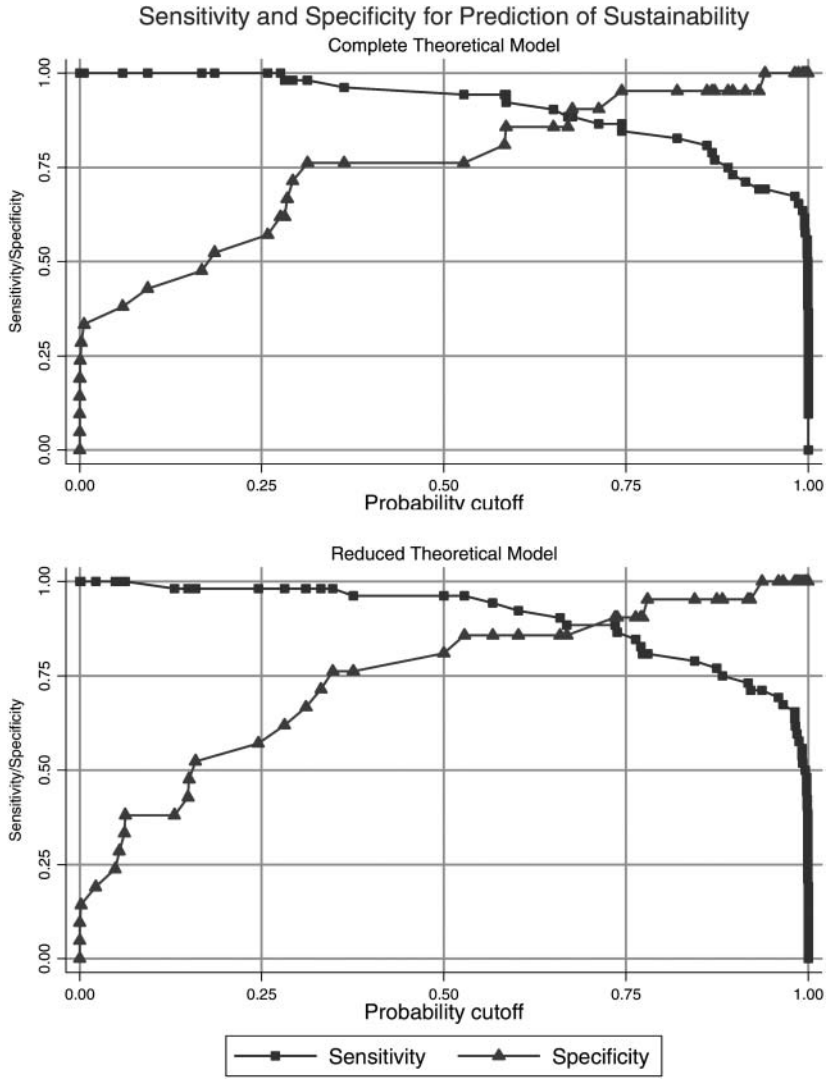


***Figure 3.*** Sensitivity and specificity of predictions of sustainability at different probability cutoffs for 74 teachers who participated in the follow-up study, based on complete theoretical model (top) and reduced theoretical model (bottom).

claiming a whole language perspective predicted *lower* likelihood of sustaining KPALS; the other variables predicted a higher level of sustainability.

For our third question (Did the level of external technical support provided by our research team moderate the likelihood of sustainability?), we found unexpected results. Membership in the KPALS-W+B condition was not

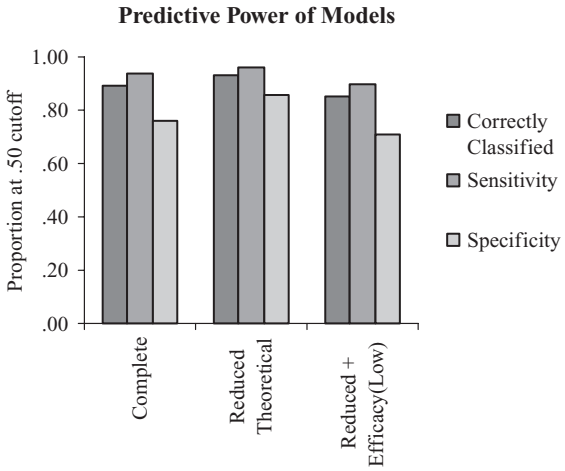


***Figure 4.*** Proportion of correct classifications, sensitivity, and specificity of the three model types.

a statistically significant predictor of sustainability relative to the KPALS-W condition ($-2.63$ [$SE = 1.69$], $p = .12$). Participation in the KPALS-W+B+H condition (present only at the Tennessee site), on the other hand, was a statistically significant predictor of a *lower* likelihood of sustaining ($-12.76$ [$SE = 4.61$], $t = -2.77$) compared to KPALS Workshop condition. Testing the difference between KPALS-W+B and KPALS-W+B+H showed that the latter condition was significantly less likely to sustain ($\chi^2 = 8.32$, $p = .004$).

After examining the individual coefficients for the variables, we calculated overall probabilities of sustaining and confidence intervals for outcomes relevant to our research questions. We examined the effects of perceived effectiveness and treatment condition on the probability of sustaining, in combination with other variables. These probabilities and confidence intervals are shown in Figures 5 and 6. We found that teachers with effectiveness ratings at least 1 *SD* away from the mean had significantly different probabilities of sustaining than teachers at the mean. We also found that the probability of sustaining for a teacher in the KPALS-W+B+H condition working in Tennessee (there were KPALS-W+B+H teachers only at the Tennessee site), was significantly lower than for teachers in the KPALS-W and the KPALS-W+B conditions across all sites.

## Reduced Theoretical Model With Effectiveness for Low-Achieving Students Only

For the final question (Is the perceived effectiveness of KPALS for all students or the perceived effectiveness of KPALS for low-achieving students a

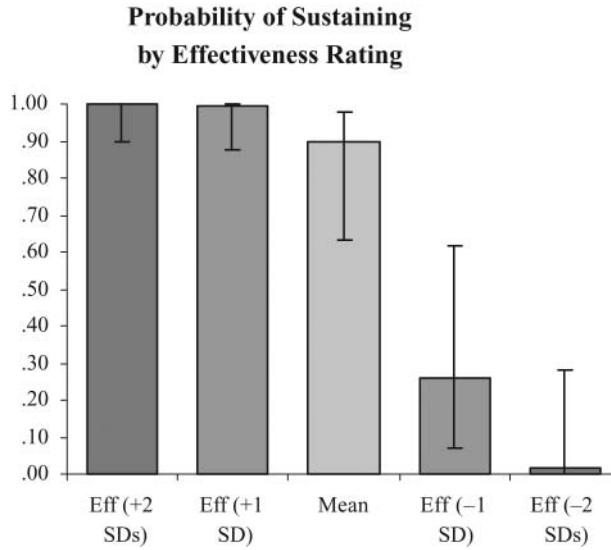


***Figure 5.*** Probabilities of sustaining Kindergarten Peer-Assisted Learning Strategies for teachers with different ratings of the effectiveness of the program for their students.

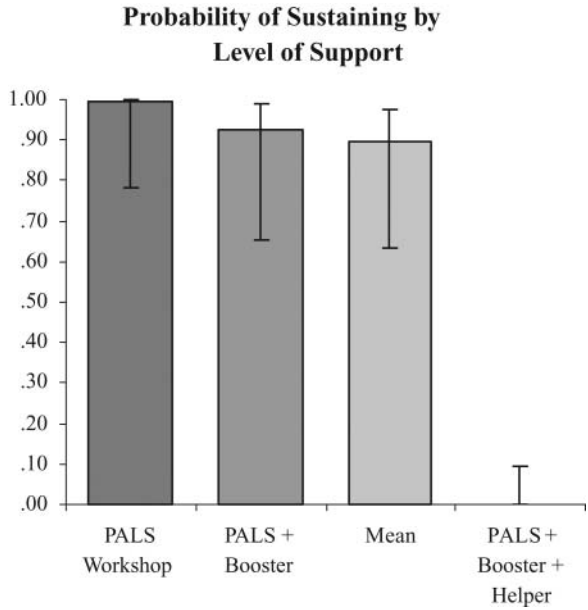


***Figure 6.*** Probabilities of sustaining Kindergarten Peer-Assisted Learning Strategies for teachers in different treatment conditions.

better predictor of sustainability?), we found that replacing overall perceived effectiveness with perceived effectiveness for low-achieving students did *not* produce a more accurate model. Pearson and Hosmer-Lemeshow chi-square statistics were larger, although not statistically significant. Predictive accuracy was also lower (see Figure 4). In addition, many coefficients, including that for perceived effectiveness for low-achieving students, failed to achieve significance. We determined, therefore, that effectiveness for low students was not a greater driving force in the decision to sustain than was effectiveness for all students, contrary to the suggestion of Gersten et al. (2000).

## DISCUSSION

If research is to be sustained in practice, teachers must support evidence-based programs and sustain their use over time. Identifying what causes teachers to sustain or discontinue research-based programs is therefore very important. We believe this study contributes to that effort in three ways. First, the heuristic model of sustainability shown in Figure 2 may provide a helpful framework for future examinations of sustainability. We recognize that it may not capture all nuances of the phenomenon, but it is inclusive of many possible influences on teacher decision making and therefore a potentially useful starting point. Second, the use of logistic regression to examine predictors of sustainability is relatively new (Taylor, 2006), and we extend this approach by building a model consistent with sustainability theory. Finally, our findings are consonant with current theories about sustainability mechanisms and contribute new ideas about the forces involved in teachers' decision making.

We found that the theoretical model of sustainability was effective in predicting whether teachers did or did not sustain KPALS. A reduced version of the same model was also effective, arguably more so. The strongest positive predictor of sustainability in the reduced model was teacher perception of KPALS effectiveness. Participation in the KPALS-W+B+H condition was a strongly *negative* predictor.

### Perception of Effectiveness

Resources available to school districts for encouraging sustainability are often limited. So it is important to know not just which ones work, but which ones work best. Perceived program effectiveness was one of our strongest positive predictors, and this makes sense: Teachers will continue to use programs if they think they work. Perceived program effectiveness for predicting sustainability may become even more important if standards-based policy continues to emphasize academic achievement. Improving sustainability may, therefore, be

partly a matter of showing teachers evidence that their instructional programs have promoted their students' academic growth.

In the case of KPALS, however, average student gains do not appear to correlate strongly with perceived program effectiveness. That is, teachers whose students demonstrate the greatest gains do not view PALS more positively than teachers whose students earn the most modest gains. We presume teachers' perceptions are not capricious, so we suggest the possibility that their perceptions are shaped in part by external technical support, that their interactions with our project staff may have caused many of them to view KPALS favorably. In general, then, providers of external technical support may be able to shape teachers' perceptions of a program's effectiveness, apart from its actual influence on student achievement. This possibility could be dangerous in the hands of zealous promoters, especially because schools can cling to practices with few or no effects (Malouf & Schiller, 1995). On the other hand, the importance of perceived program effectiveness may help providers of external support when programs may be effective, but not immediately so. Because reform efforts sometimes take multiple years to demonstrate achievement effects (e.g., Berends et al., 2002; Desimone, 2003), providers of external technical support may increase early implementation and later sustainability by drawing teachers' attention to small initial improvements in student performance.

## External Technical Support

Our examination of the level of external technical support (i.e., the participation in the KPALS-W, KPALS-W+B, or KPALS-W+B+H conditions) yielded an unexpected result: When we gave teachers more support, it *reduced* the likelihood they would sustain KPALS. This was surprising because it runs counter to prior research (e.g., Han & Weiss, 2005; McLaughlin, 1990) and because greater external technical support did improve KPALS fidelity (Stein et al., 2008) and student achievement (Fuchs et al., 2008).

We propose that the negative influence of participation in the KPALS-W+B+H condition on sustainability argues for the importance of flexibility in providing external technical support, as Datnow et al. (1998) suggested. Helpers in our study were instructed to work with teachers to improve their fidelity of implementation, leaving little room for them to adapt the program to their own circumstances. An inadvertent consequence may have been a kind of "KPALS fatigue," where teachers felt confined to continue the program as designed. A second unintended consequence of intensive technical support may have been dependency on the Helper. In other words, teachers may have become so accustomed to the additional support that they never achieved anything close to independent implementation. This, too, may have reduced a sense of program ownership and led to a decision to discontinue its use.

By contrast, teachers in the KPALS-W and KPALS-W+B conditions had de facto latitude (i.e., we did not encourage departure from expected implementation, but the absence of a Helper may have produced it) to use KPALS in a manner that suited them personally and connected with the needs of their students. Their KPALS may have become just that: *Theirs*. By personalizing the program, they may have made it their own and in so doing increased their investment in it. This possibility is consonant with the idea that external technical support should emphasize the importance of modest forms of modifiability, or customization, of evidence-based programs by teachers if they are to be sustained by them.

Given that there are benefits to compliance-focused fidelity—our prior work showed it improves adherence and student achievement—we suggest that overemphasizing the personalizing or customizing of an educational intervention could damage its integrity. We believe there is a (still unchartered) middle ground, a compromise to be found: Providers of external technical assistance emphasize fidelity to *core elements* and permit customization of elements they consider less central to program success. This might contribute to greater sustainability but assure the key elements of the program remain in place. Such an approach should be the focus of systematic research.

## LIMITATIONS

Although the complete theoretical model predicted sustainability relatively well, there are reasons to be skeptical. First, the use of 15 predictors in the complete model and 12 in the reduced raises the concern that we overfitted the model (Peduzzi et al., 1996). Simulation studies have raised the concern that when the smaller number of observed events (here, 21, the number of nonsustaining teachers) is less than 10 times the number of predictors, significant effects may not generalize well (Peduzzi et al., 1996; Steyerberg et al., 2001). Second, we have several reasons to suspect omitted variable bias. One, we could not model two theoretically important dimensions of sustainability (training quality and feasibility). Two, the significant effect for the Minnesota site suggests some factor specific to that site not otherwise captured in the model. Three, we did not have sufficient power to examine interactions between variables. An option for correcting this bias, instrumental variables, was not viable because of our relatively small sample size and the absence of valid instruments.

Another obvious limitation is that we used teacher self-report data. The finding that teachers who worked with Helpers were less likely to sustain their use of KPALS may reflect a conservative understanding of what constitutes "use." Having had regular discussions with Helpers about the correct way to implement KPALS, they may have regarded their highly modified use as nonuse. By contrast, their teacher counterparts without Helper support may

have applied a less restrictive definition of use. In addition, the teacher self-report data were collected by means of interviews and surveys. There are, of course, different ways of measuring the variables in our analysis, and collecting data in different ways might have changed our results.

Moreover, although we hope this study contributes to future work on sustainability, we caution readers that the results may be unique to a KPALS scaling-up context. The sample was self-selected and statistically significantly different from noninterviewed KPALS teachers on two variables used in the reduced model (perceived effectiveness and fidelity). Furthermore, teachers who participated in the KPALS study opted to do so. Given that teachers who choose to participate in reforms tend to differ from those who do not (Bodilly, Keltner, Purnell, Reichardt, & Schuyler, 1998), our sample may be different than teachers in general. Given the substantive differences between KPALS and more comprehensive schoolwide reforms, our predictors of teacher sustainability may differ to a great extent. In addition, sustaining a program in Kindergarten is different than sustaining a program in higher grades, where the emphasis on test preparation is greater (Datnow & Stringfield, 2000) and the strength of implementation declines (Bodilly et al., 1998). Our ability to generalize is also no doubt affected by the fact that this study examines "sustainability" only 1 year after external technical support ended. As Coburn (2003) pointed out, reform is really sustained when teachers have used it independently for several years. Factors that predicted sustained use of KPALS after 1 year may differ from those predicting its use after 5.

As described at the outset, larger numbers of evidence-based practices are emerging from rigorous experimental research and the current policy climate favors their use. The presence of incentives for using such programs does not guarantee their long-term use, however. It remains critical that program developers, policymakers, and school officials better understand how to convince teachers and schools to continue using instructional programs that have positive outcomes for their students. The field has done much to develop our understanding of these issues; it remains for us to refine our understanding of the critical levers in sustainability and engage in further research that will test their importance. We see our study as one step in that direction.

## REFERENCES

Ball, D. L. (1995). Blurring the boundaries of research and practice. *Remedial and Special Education*, *16*, 354–363.

Berends, M., Chun, J., Schuyler, G., Stockly, S., & Briggs, R. J. (2002). Implementation of NAS designs in a high-poverty district. In M. Berends, S. J. Bodilly, & S. N. Kirby (Eds.), *Facing the challenges of whole-school reform: New American Schools after a decade* (pp. 95–122). Santa Monica, CA: RAND Corporation.

Boardman, A., Arguelles, M. E., Vaughn, S., Hughes, M. T., & Klingner, J. (2005). Special education teachers' views of research-based practices. *The Journal of Special Education*, *39*, 168–180.

Bodilly, S. J., Keltner, B., Purnell, S., Reichardt, R., & Schuyler, G. (1998). *Lessons from New American Schools' scale-up phase: Prospects for bringing designs to multiple schools*. Santa Monica, CA: RAND Corporation.

Borman, G. D., Slavin, R. E., Cheung, A., Chamberlain, A. M., Madden, N. A., & Chambers, B. (2006). Success For All: First-year results from the national randomized field trial. *Educational Evaluation and Policy Analysis*, *27*, 1–22.

Bos, C., Mather, N., Dickson, S., Podhajski, B., & Chard, D. (2001). Perceptions and knowledge of preservice and inservice educators about early reading instruction. *Annals of Dyslexia*, *51*, 97–120.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.

Carnine, D. W. (1997). Bridging the research-to-practice gap. In J. W. Lloyd, E. J. Kameenui, & D. Chard (Eds.), *Issues in educating students with disabilities* (pp. 363–376). Mahwah, NJ: Erlbaum.

Chall, J. S. (2000). *The academic achievement challenge: What really works in the classroom?* New York: Guilford.

Coburn, C. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher*, *32*(6), 3–12.

Darling-Hammond, L. (1996). The quiet revolution: Rethinking teacher development. *Educational Leadership*, *53*(6), 4–10.

Datnow, A. (2002). Can we transplant educational reform, and does it last? *Journal of Educational Change*, *3*, 215–239.

Datnow, A. (2005). The sustainability of comprehensive school reform models in changing district and state contexts. *Educational Administration Quarterly*, *41*, 121–153.

Datnow, A., & Castellano, M. (2000). Teachers' responses to Success For All: How beliefs, experiences, and adaptations shape implementation. *American Educational Research Journal*, *3*, 775–799.

Datnow, A., Hubbard, L., & Mehan, H. (1998). *A framework for understanding educational reform implementation*. Santa Cruz, CA: Center for Research on Education, Diversity, & Excellence.

Datnow, A., & Stringfield, S. (2000). Working together for reliable school reform. *Journal of Education for Students Placed At Risk*, *5*, 183–204.

Desimone, L. (2003). How can comprehensive school reform models be successfully implemented? *Review of Educational Research*, *72*, 433–479.

Dusenbury, L., Brannigan, R., Falco, M., & Hansen, W. B. (2003). A review of research on fidelity of implementation: Implications for drug abuse prevention in school settings. *Health Education Research*, *18*, 237–256.

Elmore, R. (1996). Getting to scale with good educational practice. *Harvard Educational Review*, *66*(1), 1–26.

Fuchs, D., & Fuchs, L. S. (1998). Researchers and teachers working together to adapt instruction for diverse learners. *Learning Disabilities Research & Practice*, *13*, 126–137.

Fuchs, D., & Fuchs, L. S. (2005). Peer-Assisted Learning Strategies: Promoting word recognition, fluency, and reading comprehension in young children. *Journal of Special Education*, *39*(1), 34–44.

Fuchs, D., Fuchs, L. S., Al Otaiba, S., Thompson, A., Yen, L., McMaster, K. N., et al. (2001). KPALS: Helping Kindergarteners with reading readiness: Teachers and researchers in partnerships. *Teaching Exceptional Children*, *33*, 76–80.

Fuchs, D., Fuchs, L., Mathes, P., & Simmons, D. (1997). Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, *34*, 174–206.

Fuchs, D., Fuchs, L, Thompson, A., Al Otaiba, S., Yen, L., Yang, N., et al. (2001). Is reading important in reading-readiness programs? A randomized field trial with teachers as program implementers. *Journal of Educational Psychology*, *93*, 251–267.

Fuchs, D., Fuchs, L. S., Thompson, A., Svenson, E., Yen, L., Yang, N., et al. (2001). Peer-assisted learning strategies in reading: Extensions for kindergarten, first grade, and high school. *Remedial and Special Education*, *22*, 15–21.

Fuchs, D., McMaster, K., Saenz, L., Yen, L., Taylor, K., Lemons, C., et al. (2008). Scaling-up Kindergarten PALS: A multi-site randomized control trial to explore how much teacher support is necessary. In D. Fuchs (Chair), *Scaling-up Kindergarten Peer-Assisted Learning Strategies and students' reading achievement: A multi-site longitudinal randomized control trial*. Annual Convention of the American Educational Research Association. New York City.

Gersten, R., Baker, S. K., Smith-Johnson, J., Flojo, J. R., & Hagan-Burke, S. (2004). A tale of two decades: Trends in support for federally funded experimental research in special education. *Exceptional Children*, *70*, 293–304.

Gersten, R., Chard, D., & Baker, S. (2000). Factors enhancing sustained use of research-based instructional practices. *Journal of Learning Disabilities*, *33*, 445–457.

Gersten, R., & Dimino, J. (2001). The realities of translating research into classroom practice. *Learning Disabilities Research & Practice*, *16*, 120–130.

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149–164.

Gersten, R., Vaughn, S., Deshler, D., & Schiller, E. (1997). What we know about using research findings: Implications for improving special education practice. *Journal of Learning Disabilities*, *30*, 466–476.

Giles, C. (2006). Sustaining secondary school visions over time: Resistance, resilience and educational reform. *Journal of Educational Change*, *7*, 179–208.

Glennan, T., Bodilly, S. J., Galegher, J. R., & Kerr, K. A. (2004). *Expanding the reach of educational reforms: Perspectives from leaders in the scale-up of educational interventions*. Santa Monica, CA: RAND Corporation.

Greenwood, C. R., & Abbott, M. (2001). The research to practice gap in special education. *Teacher Education and Special Education*, *24*, 276–289.

Greenwood, C. R., Delquadri, J. C., & Hall, R. V. (1989). Longitudinal effects of Classwide Peer Tutoring. *Journal of Educational Psychology*, *81*, 371–383.

Han, S. S., & Weiss, B. (2005). Sustainability of teacher implementaiton of school-based mental health programs. *Journal of Abnormal Child Psychology*, *33*, 665–679.

Hargreaves, A., & Goodson, I. (2006). Educational change over time? The sustainability and nonsustainability of three decades of secondary school change and continuity. *Educational Administration Quarterly*, *42*, 3–41.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley & Sons, Inc.

Klingner, J. K., Ahwee, S., Piloneta, P., & Menendez, R. (2003). Barriers and facilitators in scaling up research-based practices. *Exceptional Children*, *69*, 411–429.

Malouf, D. B., & Schiller, E. P. (1995). Practice and research in special education. *Exceptional Children*, *61*, 414–424.

McLaughlin, M. W. (1990). The Rand Change Agent Study revisited: Macro perspectives and micro realities. *Educational Researcher*, *19*(9), 11–16.

McLaughlin, M. W., & Mitra, D. (2001). Theory-based change and change-based theory: Going deeper, going broader. *Journal of Educational Change*, *2*, 301–323.

McMaster, K. L., Fuchs, D., & Fuchs, L. S. (2006). Research on Peer-Assisted Learning Strategies: The promise and limitations of peer-mediated instruction. *Reading and Writing Quarterly*, *22*, 5–25.

McMaster, K. L., Kung, H., Han, I., & Cao, M. (2007). Peer-Assisted Learning Strategies: A "tier 1" approach to promoting English learners' response to intervention. *Exceptional Children*, *74*, 194–214.

No Child Left Behind Act, PL 107-110 (2001).

O'Connor, R. E. (1999). Teachers learning Ladders to Literacy. *Learning Disabilities Research & Practice*, *14*, 203–214.

Peduzzi, P., Concato, J., Kemper, E., Holford, T. R., & Feinstein, A. R. (1996). A simulation study of the number of events per variable in logistic regression analysis. *Journal of Clinical Epidemiology*, *49*, 1373–1379.

Pressley, M. (2006). *Reading instruction that works: The case for balanced instruction* (3rd ed.). New York: Guilford.

Rand Corporation (1978). Volume VIII: Implementing and sustaining innovations. (Research Rep. No. R-1589/8 HEW). Santa Monica, CA: Author.

Rohrbach, L. A., Graham, J. W., & Hansen, W. B. (1993). Diffusion of a school-based substance abuse prevention program: Predictors of program implementation. *Preventive Medicine*, *22*, 237–260.

Rose, D. J., & Church, R. J. (1998). Learning to teach: The acquisition and maintenance of teaching skills. *Journal of Behavioral Education*, *8*, 5–35.

Sáenz, L. M., Fuchs, L. S., & Fuchs, D. (2005). Peer-Assisted Learning Strategies for English language learners with learning disabilities. *Exceptional Children*, *71*, 231–247.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton Mifflin.

Slavin, R. E., Madden, N. A., & Datnow, A. (2007). Research in, research out: The role of research in the development and scale-up of Success for All. In S. Fuhrman, D. Cohen, & F. Mosher (Eds.). *The state of education policy research* (pp. 261–280). Mahwah, NJ: Erlbaum.

Slavin, R. E., Madden, N. A., Dolan, L. J., Wasik, B. A., Ross, S., Smith, L., et al. (1996). Success For All: A summary of research. *Journal of Education for Students Placed at Risk*, *1*, 41–76.

Stein, M. L., Berends, M., Fuchs, D., McMaster, K., Sáenz, L., & Yen, L. (2008). Factors related to fidelity of a peer-assisted learning strategies intervention for kindergarten reading. *Educational Evaluation and Policy Analysis*, *30*, 268–388.

Steyerberg, E. W., Harrell, F. E., Borsboom, G. J. J. M., Eijkemans, M. J. C., Vergouwe, Y., & Habbema, J. D. F. (2001). Internal validation of predictive models:

Efficiency of some procedures for logistic regression analysis. *Journal of Clinical Epidemiology*, *54*, 774–781.

Stringfield, S., Datnow, A., & Ross, S. M. (1998). *Scaling up school restructuring in multicultural, multilingual contexts: Early observations from Sunland County*. Berkeley, CA: Center for Research on Education, Diversity, & Excellence.

Taylor, J. E. (2006). The struggle to survive: Examining the sustainability of schools' comprehensive school reform efforts. *Journal of Education for Students Placed At Risk*, *11*, 331–352.

Tyack, D., & Tobin, W. (1994). The "grammar" of schooling: Why has it been so hard to change? *American Educational Research Association*, *31*, 453–479.

Vadasy, P. F., Jenkins, J. R., Antil, L. R., Phillips, N. B., & Pool, K. (1997). The research-to-practice ball game: Classwide peer tutoring and teacher interest, implementation, and modifications. *Remedial and Special Education*, *18*, 143–156.

Vaughn, S., Hughes, M. T., Schumm, J. S., & Klingner, J. (1998). A collaborative effort to enhance reading and writing instruction in inclusion classrooms. *Learning Disability Quarterly*, *21*, 57–74.

Whitehurst, G. J. (2003). *The Institute of Education Sciences: New wine, new bottles*. Paper presented at the Annual Meeting of the American Educational Research Association.