



**HOME WORKS! THE TEACHER HOME VISIT
PROGRAM: EXPLORING OUTCOMES DURING THE
2014-15 SCHOOL YEAR**

SEPTEMBER 29, 2016

SUBMITTED TO:

HOME WORKS! The Teacher Home Visit
Program
225 Linden Avenue
St. Louis, MO 63105

SUBMITTED BY:

Concentric Research & Evaluation
6 Trowbridge Lane
Mansfield, MA 02048
Lauren.scher@concentric-cre.com

In Partnership with:
EMT Associates, Inc.
1631 Creekside Drive, Suite 100
Folsom, CA 95630

TABLE OF CONTENTS

HOME WORKS! OUTCOME ANALYSIS:	1
FINDINGS AT A GLANCE	1
RESEARCH SUMMARY	2
Using lessons learned from the implementation report to help frame the outcome study	2
How did we analyze student outcomes?	2
HW! participants scored higher on the STAR assessment than similar non-HW! students.....	3
HW! participants missed fewer days of school than similar non-HW! students	4
HW! participants were equally likely to have disciplinary referrals as similar non-HW! students.....	5
Conclusions and next steps	6
APPENDIX A	8

HOME WORKS! OUTCOME ANALYSIS: FINDINGS AT A GLANCE

STUDY FACTS:

- This study examines whether students participating in HOME WORKS! The Teacher Home Visit Program (HW!) scored higher on standardized tests, had higher attendance rates, and had fewer reported disciplinary incidences than similar students who did not participate in HW! during the 2014-15 school year.
- It examines school records for up to nearly 3,000 students attending four school districts in and around the St. Louis region.
- HW! participants are compared to similar non-HW! students in the same school. They are also compared to similar non-HW! students attending matched schools

STUDY FINDINGS:

- When comparing HW! participants to similar students in the same school, those who received at least one home visit:
 - ➔ scored 5 percent higher on the STAR literacy assessment
 - ➔ were 13 percent less likely to miss at least two weeks of school
 - ➔ had similar levels of disciplinary referrals as non-HW! students.
- When comparing HW! participants to similar students in the same school, those who participated in at least two home visits:
 - ➔ scored nearly 7 percent higher on the STAR literacy assessment
 - ➔ were 25 percent less likely to miss at least two weeks of school
 - ➔ had similar levels of disciplinary referrals as non-HW! students.
- When comparing HW! participants to similar students attending a different school:
 - ➔ there were no significant differences between the groups for any outcomes.

WHAT IS NEXT?

- HW! continues to improve and strengthen its program model and promote strong implementation, which may yield stronger effects in the future.
- The within-school analysis reports promising results, but the research design limits our confidence that HW! *caused* these student improvements (and the magnitude of the improvements). A rigorous study designed to enable direct comparison of students in classrooms selected at random to implement HW! or not would yield credible evidence of program effectiveness for educators and funders.

RESEARCH SUMMARY

This report examines whether students who participate in HOME WORKS! The Teacher Home Visit Program (HW!) score higher on standardized test scores, have higher attendance rates, and have fewer disciplinary incidents than non-HW! students. It is a companion to *HOME WORKS! The Teacher Home Visit Program 2014-15 Implementation Evaluation Report*, which discusses program participation rates, program implementation, and parent and teacher perceptions of the program. In this report, we present evidence to suggest that HW! participants have stronger academic outcomes than similar students who did not participate in HW!. This summary describes the study sample, methods, results, and implications of these findings as well as suggestions for next steps. Appendix A provides a more detailed and technical discussion of the research methods.

Using lessons learned from the implementation report to help frame the outcome study

During the 2014-15 school year, the HW! organization, in partnership with EMT Associates Inc., collected a wide variety of data to explore how the HW! program is being implemented, to learn more about how teachers and parents experience the program, and to examine whether students participating in the program have improved academic outcomes. The implementation report documents broad implementation of the program across a variety of schools in multiple school districts. It summarizes teacher and student participation rates, which vary from school-to-school and within schools. The report notes, for example, that many targeted families participated in one home visit, but very few families participated in the full program model (including two home visits and two family dinners). While teachers and parents report challenges they faced in fully implementing the program in terms of completing two home visits and family dinner attendance, most parents and teachers who participated in home visits believe that HW! improves parent engagement and the quality of parent and teacher relationships. Teachers overwhelmingly report that they observed improvement in student classroom behavior and academic performance. The outcomes analysis that follows begins to explore whether there is evidence, based on student administrative academic and behavioral records, that supports these teacher-reported observations.

How did we analyze student outcomes?

The student outcomes analysis examined whether students who participated in HW! had higher test scores, improved attendance, and fewer reported disciplinary incidents than similar students who did not participate in HW!. The evaluation team examined up to 2,744 school records from the 2014-15 school year for both HW! and non-HW! students attending four school districts in and around St. Louis.¹

About HOME WORKS! The Teacher Home Visit Program (HW!)

HW! is a teacher home visiting intervention aimed at bringing teachers and families together as partners to support students' academic success. The HW! program model examined in this report includes teacher participation in two training sessions, up to two home visits per student per academic year (two teachers attend each home visit), and attendance at two annual Family Dinners at the school. Two teachers per school act as site coordinators to help support teachers who participate in the program.

¹ Student records included in this analysis had to have all required outcome information, and no missing information related to grade level, race/ethnicity or gender. See Appendix A for more details concerning sample sizes for each specific analysis.

We compared HW! students to non-HW! students in two main ways. First, we compared HW! students to other students who attended the same school but did not receive any home visits. This is called the “within-school” analysis. This analysis is useful because it compared students in the same educational setting who were experiencing the same school curriculum and other programming. We matched each HW! student with a similar student who did not participate in HW!. The downside to this comparison is that families who agree to home visits may be different in important ways from those who do not choose to participate (or who are not targeted for participation). Thus, even though we were able to find matched students who looked similar in terms of demographics and test scores, they may have been different in important ways that we could measure.²

For this reason, we also conducted a second set of comparisons called the “between-schools” analysis to see if the study results would be similar when using a different set of matched students. Each of the 10 HW! schools included in this study was paired with 10 different comparison schools in the same district. Comparison schools were chosen based on serving roughly similar kinds of students in terms of overall school demographics. In this analysis, we matched each student attending a HW! school and participating in HW! with a similar student attending a different school (and thus would never have had the opportunity to participate in HW!). We used matching strategies to ensure that students were similar in terms of prior test scores, race, gender and grade level. The downside to this comparison is that, particularly in the smaller districts, despite our attempt to find similar matched schools, the students served in these schools were quite different and school policies and practices likely varied. Similar to the within-school analysis, we were only able to match students on a very limited set of characteristics available through administrative records, thus reducing our confidence that we are comparing equivalent groups of students. We discuss this challenge in more detail in the summary section below.

Because students experience HW! in different ways (as described in the implementation study), we examined HW! participation in two different ways. First, we examined the effects for all students who received at least one home visit. Second, we focused on the smaller subset of students who took part in both home visits.³

HW! participants scored higher on the STAR assessment than similar non-HW! students

The STAR literacy assessment is a commonly used standardized assessment for elementary school youth and is administered in all four participating school districts. HW! students scored higher on the Spring 2015 STAR assessment than similar students who did not receive any home visits, but attended the same school (Figure 1). Specifically,

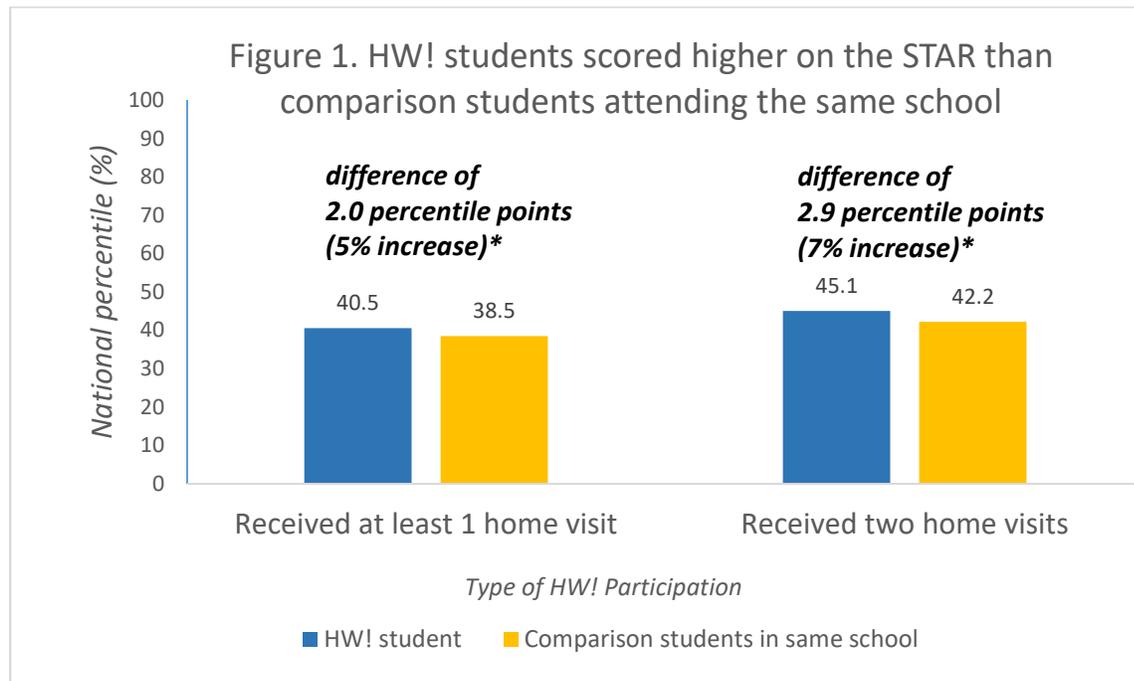
- HW! participants who received at least one home visit scored **5 percent higher** than comparable students who attended the same school and did not receive home visits. This corresponds to a 2-percentile point difference on the STAR achievement test with HW! students averaging in the 40.5th percentile and comparison students scoring, on average, in the 38.5th percentile.

² See Appendix A for more details concerning the matching procedure and the reason why a 1:1 match was necessary for these analyses.

³ Because teachers may change their practices after having taken part in home visits, we also explored whether simply being enrolled in HW! teacher’s classroom (regardless of whether a student’s family participated in home visits) leads to academic improvement. We did not find any evidence to suggest that this is occurring and, thus for simplicity, these results are not discussed in this report.

- HW! participants who took part in two home visits scored nearly **13 percent higher** than comparable students who did not receive any home visits. This corresponds to a 3-percentile point difference on the STAR test between HW! participants, who scored in the 45th percentile, and comparison students, who scored, on average, in the 42nd percentile.

Although the within-school analysis yielded results that researchers consider “statistically significant” (meaning that we can say with 95% confidence that these differences between the HW! participants and comparison group were not due to random chance), the alternate analysis that compared HW! participants to similar students attending a different school yielded no significant differences. For the between-school analysis, the average STAR scores were nearly identical for HW! participants and their matched group of students.



* All differences reported in this figure are statistically significant (p<.05). HW! participant averages presented in this analysis are regression-adjusted to account for pre-existing differences in prior STAR scores, as well as demographic differences. See Appendix A for more details.

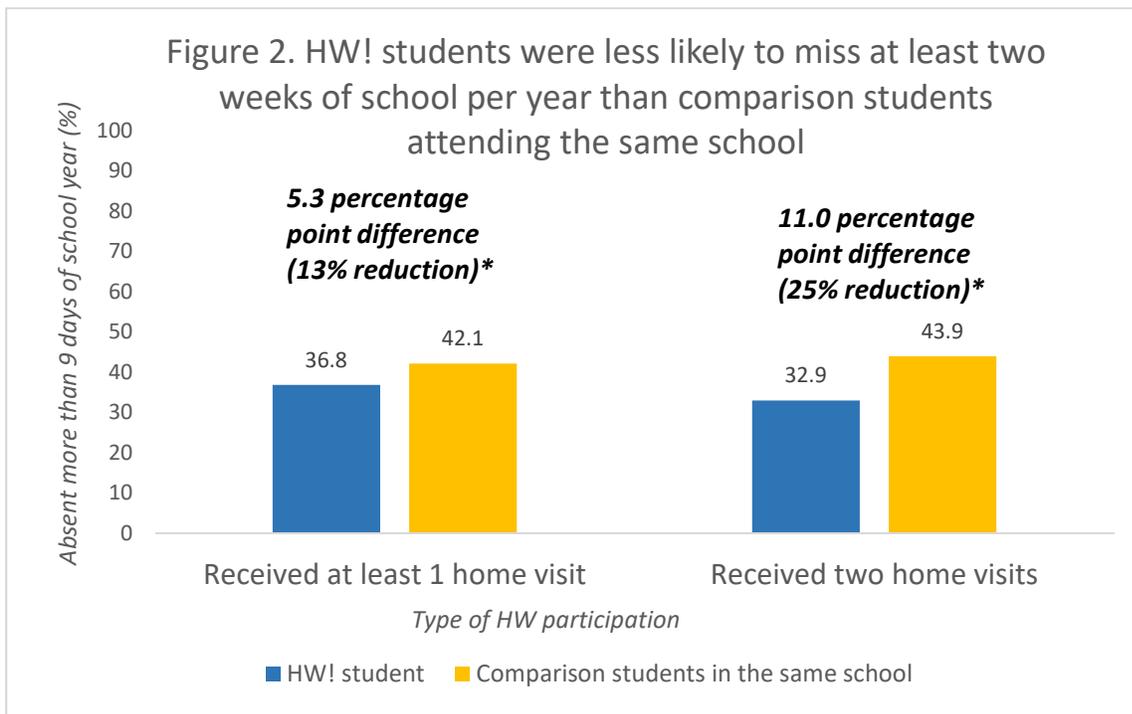
HW! participants missed fewer days of school than similar non-HW! students

We measured attendance in a variety of ways, but focused our analysis on a measure that appeared to be a major cutoff point with this particular sample of youth, missing more than 5 percent of the school year versus 5 percent or fewer days.⁴ This corresponds to missing at least two weeks of school (more than 9 days of school) in a school year, which is a meaningful amount of time to miss school for elementary school-aged youth. For the within-school analysis, HW! participants were significantly less likely to miss at least two weeks of school compared to similar non-HW! students. Specifically,

⁴ We did not observe any differences in chronic absenteeism, which, in general is uncommon in the early grade levels. On average, 11 percent of the sample would be defined as being chronically absent based on the commonly accepted definition of being absent for more than 10 percent of the school year (approximately 18 days).

- HW! participants who participated in one visit had nearly **7 percent lower rates** of missing at least two weeks of school than comparable non-HW! students. Forty-two percent of non-HW! students missed at least two weeks of school during the 2014-15 school year compared to an estimated 37 percent who took part in at least one home visit, a statistically significant five percentage point difference.
- HW! students who received at least two home visits had **25 percent lower rates** of missing at least two weeks of school than comparable non-HW! students. On average, 44 percent of non-HW! comparison students missed more two weeks of school versus 33 percent of HW! students who took part in two home visits, an 11 percentage point difference.

Similar to the STAR analysis, the between-school analysis yielded no statistically significant differences in attendance. Attendance rates for the HW! participants who received one visit versus comparison students attending other schools are nearly identical. However, if we relax the statistical standard in the spirit of an exploratory investigation to 90 percent confident ($p < .10$), we find a 5.4 percentage point difference in the rate of students missing at least two weeks of school between HW! students who took part in two home visits (37.5 percent missed more than 2 weeks of school) versus comparable students attending non-HW! schools (42.9 percent).



* All differences reported in this figure are statistically significant ($p < .05$). HW! participant averages presented in this analysis are regression-adjusted to account for pre-existing differences in prior attendance rates, as well as demographic differences. See Appendix A for more details.

HW! participants were equally likely to have disciplinary referrals as similar non-HW! students

On average, 10 percent of students for whom we have discipline behavior data available had at least one reported disciplinary incident during the 2014-15 school year. This varied across districts from as

low as 7% to a high of 18% for the sample of students we obtained. The number of disciplinary reports per student ranged from zero to 13. However, very few students had more than one disciplinary incident. Because districts may have had very different policies regarding disciplinary referrals and because it was rare to report more than one incident, we focused our analysis on whether or not a student had at least one disciplinary incident.

Regardless of whether we compared students within schools or we compared HW! participants to non-HW! students at different schools, we found no statistically significant differences in reports of behavioral incidents.⁵ Because these formal disciplinary reports are infrequent in the early grade levels, it is not the strongest measure to use to assess changes in student behaviors. Analyses of HW! impacts on disciplinary referrals at the middle or high school level may potentially yield stronger results.

Conclusions and next steps

This report demonstrates that students who participate in HW! have higher standardized literacy scores, are less likely to miss at least two weeks of school during the year, and have similar levels of disciplinary incidences than similar students attending the same schools. While these results are promising, they should be interpreted cautiously.

Because of the research design employed, it is not possible to state, for certain, that HW! *causes* academic improvement. As was discussed in the implementation report, students who participate in home visits often differ from those who do not in terms of background characteristics (such as race/ethnicity) and prior student achievement levels. In this report, we use matching methods that try to account for these pre-existing differences in the characteristics of HW! students and non-HW! students. However, it is quite likely that there are differences between families who choose to participate in HW! and those who do not in ways that we cannot measure. For example, a family who is interested in taking part in home visits may also be more motivated to support their child's academic progress. This motivation may mean that families who take part in home visits may already have more books in the home or are working more closely with their children than families who are not willing to take part in home visits. This may help explain differences in outcomes. It is also possible that the opposite is the case, and there would have been even stronger effects had more equivalent groups been compared. For example, in classrooms where teachers prioritized students who they perceived to be the most at risk, it is likely that the HW! participant families had fewer resources in the home to support learning than non-HW! families.

The fact that the between-school analysis, where HW! participants were compared with similar students attending other schools, yielded no significant differences on academic outcomes further points out that the promising results presented in this report should be considered with caution. The between-school analysis, however, is not without its own flaws. In addition to the same concerns that it was only possible to match students on a limited set of pre-existing characteristics, an added concern with this between-school comparison is that school policies and other school-level programming may also explain

⁵ For the within-school analysis, there was a 1.1 percentage point different in the percentage of students with a disciplinary incident report (an average of 8.6 for HW! participants who had one home visit versus 9.7 for non-HW! students). The difference was 3.0 percentage points for the two-visit participants versus comparison analysis (7.8 for HW! participants versus 10.8 for comparison students). These differences were not statistically significant, and thus we cannot rule out whether these differences between the groups are due to chance.

differences in outcomes. For example, schools may have very different school-level policies or practices that may influence the number of disciplinary referrals. Thus, any differences, or lack of differences in this outcome may simply be related to school characteristics and unrelated to the home visiting intervention.

While the results from this study are exploratory, they do point to interesting trends for further exploration. In particular, the magnitude of differences in outcomes are larger when comparing students who took part in two home visits versus non-HW! students, as opposed to the analysis that focused on students who had only one home visit. Future analyses should continue to explore whether this pattern continues, and to explore the strength of participation in family dinners in addition to home visits.

A rigorous randomized controlled trial study (RCT) is regarded as the “gold standard” for determining program effectiveness. Such a study would select students or classrooms to participate in HW! based on a random process would enable us to have considerable confidence that any differences that we find between HW! and non-HW! students are directly due to the HW! program, not other factors.⁶ When randomization is employed, the groups being compared should be similar on characteristics that can be measured easily (such as pretest scores) as well as characteristics that cannot (such as parent motivation or family resources), with the remaining difference being access to HW!.

HW! also continues to strengthen its teacher-training procedures and refine the program to be more responsive to the needs of participating teachers and families. Over the past year, the HW! organization has engaged in teacher, administrator and parent focus groups and surveys, and has commissioned a panel of experts to review training and program materials in an effort to strengthen the program. HW! also continues to improve oversight efforts in an attempt to strengthen program implementation. As the program continues to grow and improve, we would expect that the impact of the program on student outcomes would also increase. Future analyses would be able to determine if this is the case.

⁶ This methodology does not necessitate providing HW! services to fewer families or classrooms. It merely structures how the HW! resources are assigned simultaneously to serve as many families/classrooms as possible in a manner that supports a rigorous analysis.

APPENDIX A

As discussed in the *HOME WORKS! The Teacher Home Visit Program 2014-15 Implementation Evaluation Report*, HW! students differed widely from non-HW! students with regard to demographic characteristics and student achievement. The problem with this lack of comparability is that it makes it difficult to attribute differences in academic outcomes between HW! students and non-HW! students to the HW! program. For example, if non-HW! students tended to be lower achievers than HW! students based on prior year's test scores, any observed differences between the groups in outcomes at the end of the school year may simply be due to these pre-existing differences in achievement and not anything to do with the effects of the program. Ideally, we would be able to measure program impacts using a randomized controlled trial (RCT) design, in which students or teachers are randomly selected to participate in HW! or to be in the comparison group. When randomization occurs, we would expect the intervention and comparison groups to be very similar in pre-program characteristics both in ways we can observe (for example, prior test scores) as well as in ways we cannot observe (for example, prior levels of family engagement and involvement).

Matching procedure. In the analyses presented here, students were not randomly assigned to be in the HW! program and we have clear evidence that the groups looked quite different from each other in terms of, for example, prior achievement levels and race/ethnicity. We therefore used a statistical technique called propensity score matching to identify matched pairs of students who looked similar on characteristics such as prior achievement, race/ethnicity, gender, and grade level. This quasi-experimental technique helps identify comparison groups who "look" similar to HW! students in ways that we can measure, but there are still likely many differences between these groups that we are not able to measure and that may explain differences in outcomes. Nevertheless, this approach provides a useful first step in exploring program impacts. We conducted matching both across districts for the pooled analysis and separately within district for district-specific analyses (which are not reported in this paper).⁷

Sample. The analysis uses administrative records obtained from four participating school districts. These districts included a mix of large urban and small rural systems in and around the great St. Louis region. To be included in the sample, a student would need to have had non-missing pre-program and outcome measures, as well as non-missing demographic information. The sample sizes presented in this paper ranged from 734 to 2,744, depending on the outcome and comparison sample. As discussed in the paper, we conduct this exploratory analysis with two types of comparison samples. Specifically, we first compare HW! participants to other non-HW! participants attending the *same schools*, which we refer to as the "within-school analysis." Next, we compare HW! participants to students attending other schools that do not participate in HW!. We refer to this as the "between-school analysis."

Defining HW! participation. Students could participate in the HW! program in a variety of ways. For example, students in HW! classrooms may receive up to two home visits and they may or may not also participate in up to two family dinners. For the purposes of our analysis, we define HW! participation in

⁷ In many cases, district-specific samples sizes are not large enough to have enough statistical power to detect significant effects. In addition, in some cases, we are unable to construct equivalent samples for some of the district analyses given very limited pools from which to match HW! students to non-HW! students based on the available data. We therefore exclude district-specific results from this report.

three main ways. First, we examine outcomes for students who participated in at least one home visit. This was the primary way that youth experienced HW! during the 2014-15 school year. Approximately 1,400 youth are included in this analysis. Second, we present results comparing students who received two visits to those who didn't receive any home visits, which reduces the analysis sample to just over one-half of the primary analysis (approximately 800 students).⁸

Table A.1 presents key statistics from all study results discussed in this paper. The top panel presents results for the within-school analysis, and the bottom panel focuses on the between-school results. The table presents sample sizes for each analysis, as well as baseline (end of 2013-14 school year) and regression-adjusted follow-up (end of 2014-15 school year) means and raw standard deviations for each outcome. Outcomes include, STAR literacy test scores, the percent of students missing more than 9 days of school in a year, and the percent of students who had one or more disciplinary referral during the school year. The table also presents mean differences as well as a Hedge's *g* effect size estimate, which is a standardized mean difference (in general, calculated as the mean difference divided by the pooled standard deviation).

⁸ We also examined students enrolled in HW! classrooms to students enrolled in non-HW! classrooms, which aims to measure both the direct effect of HW! combined with the indirect effect of the benefit of having a teacher who has experienced HW! and who may have adjusted their teaching behaviors because of that experience. The sample size for the analysis comparing students in HW! classrooms versus those not in HW! classrooms varies for the within school (N=734) versus the across school analysis (N=2,744). The primary reason for this difference is because, in some schools implementing HW!, there is a very small pool of non-HW! classrooms. This analysis did not yield any evidence suggesting that students in HW! classrooms have improved outcomes compared to students in non-HW! classrooms and are not discussed in the body of this paper. Comparison classrooms may not have been comparable to HW! classrooms, and thus the results are excluded from this paper.

Table A.1. Baseline, follow-up and estimated program impacts for HW! participants compared to matched samples of non-HW! students

	Baseline (end of 2013-14 school year)					Follow-up (end of 2014-15 school year)					Mean difference	Effect size ¹	Statistical significance ²
	Sample size	HW! Participants		Comparison		HW! Participants		Comparison					
		mean	stdev	mean	stdev	mean	stdev	mean	stdev				
Within-school analysis													
STAR test score													
one visit vs. no visits	1430	41.90	31.39	38.90	30.08	40.54	29.97	38.53	27.75	2.01	0.07	*	
two visits vs. no visits	794	46.32	31.82	40.58	31.54	45.07	30.37	42.20	29.65	2.87	0.10	*	
% missing at least two weeks of school													
one visit vs. no visits	1378	35.99	48.03	27.87	44.87	36.78	49.36	42.08	49.41	-5.30	-0.11	*	
two visits vs. no visits	792	30.05	45.91	25.25	43.50	32.98	47.87	43.94	49.69	-10.96	-0.22	**	
% with one or more disciplinary referrals													
one visit vs. no visits	1236	9.71	29.63	6.80	25.19	8.62	29.63	9.71	28.26	-1.09	-0.04	ns	
two visits vs. no visits	760	7.37	26.16	11.58	32.04	7.83	22.88	10.79	31.07	-2.96	-0.11	ns	
Between-school analysis													
STAR test score													
one visit vs. no visits	1432	41.84	31.41	40.71	30.68	41.57	29.99	41.22	30.34	0.35	0.01	ns	
two visits vs. no visits	784	45.91	31.78	47.08	31.80	48.67	30.38	48.49	30.35	0.18	0.01	ns	
% missing at least two weeks of school													
one visit vs. no visits	1430	34.69	47.63	35.80	47.98	44.16	0.49	43.90	49.66	0.26	0.01	ns	
two visits vs. no visits	792	30.05	45.91	32.58	46.93	37.45	0.48	42.90	49.56	-5.45	-0.16	~	
% with one or more disciplinary referrals ³													
one visit vs. no visits	1270	9.29	29.05	11.97	32.48	9.47	0.28	8.98	28.61	0.49	0.02	ns	
two visits vs. no visits	762	7.35	26.13	11.02	31.36	6.35	0.23	9.19	28.92	-2.84	-0.14	ns	

Notes: For the STAR test score, a positive number favors the HW! group and a negative number favors the comparison group. For the attendance and disciplinary referral outcomes, a negative number favors the HW! group and a positive number favors the comparison group.

¹ Effect sizes are calculated as Hedge's g standardized mean differences.

² Significance levels include *** = p<.001, ** = p<.01, * =p <.05, ~ = p<.10.

³ Disciplinary incidents analysis excludes one school district that had incorrect disciplinary information.