



AMERICAN INSTITUTES FOR RESEARCH

---

# Validation Review of Year 2 Service Study Methodology and Results

---

July 19, 2005

*Submitted to:*

Office of the Independent Monitor  
Los Angeles Unified School District

*Submitted by:*

American Institutes for Research  
Jenifer J. Harr  
Miguel Socias  
Tom Parrish

## I. Introduction and Summary of Findings

This report presents findings from a validation study conducted by AIR comparing results from its study of the delivery of special education services in the Los Angeles Unified School District (LAUSD or the “District”) for the 2003-04 school year to those derived from a similar study conducted for 2004-05 by the Program Evaluation and Research Branch (PERB) of LAUSD. The purpose of both of these studies was to measure the degree of conformity between the actual provision of special education services in LAUSD with what is required by each student’s Individualized Education Program (IEP). In short, the studies attempt to ascertain the extent to which LAUSD students are actually receiving what is required under special education law.

The federal Individuals with Disabilities Education Act (IDEA) requires that all eligible students with disabilities have an IEP, which specifies instructional goals and the special education services<sup>1</sup> for each student in special education. Students in special education may receive a broad array of services according to their IEPs. For most services, the IEP specifies the frequency and duration with which the service is to be provided. The AIR study for the 2003-04 school year developed indicators of the extent to which each of these required services was actually being provided to students with disabilities throughout the District in accordance with their IEPs. The results were used to develop a performance outcome in the area of service delivery for LAUSD as required by the Modified Consent Decree, a class action settlement requiring improvements in a number of aspects of the District’s special education system.

The PERB study for the 2004-05 school year built upon AIR’s Year 1 methodology. The purpose of this validation study is to provide a cross-check of LAUSD’s Year 2 findings in two primary ways: 1) to assess whether the 2004-05 results are an accurate depiction of service delivery in LAUSD, and 2) to review the differences in the findings between the two studies and assess to what extent the differences in results reflect genuine change in compliance rates or are a product of changes in methodology.

In both studies, there were two major research components. The first analyzed a sample of IEPs and corresponding service logs to assess the match between the types of service specified in the IEP, including duration and frequency, with what was actually provided as recorded in the provider logs. The second component involved visits to the schools of a sub-sample of students to provide a further check against service provision as recorded in logs. Observers traveled to schools at the scheduled time of service delivery to observe whether the service specified in the IEP was actually provided and for what length of time.

**Report Overview:** In the remainder of this first section, we provide an overview of findings. Subsequent sections provide a detailed account of the differences between the two studies and the potential impact on the results. Section II of the report reviews the 2003-04 results alongside the 2004-05 estimates for the overall population and by disability category (estimates by service category are in Appendix A). Comparisons are shown for the four primary analyses: evidence of

---

<sup>1</sup> Please refer to U.S. Department of Education (1999), Code of Federal Regulations (CFR) Sections 300.24 and 300.6, for descriptions of related services and assistive technology services to be made available to eligible students with disabilities.

service provision using log data, frequency compliance, duration compliance, and evidence of service provision using site visitation data.

Section III describes the differences in the methodologies between the two years and attempts to assess the impact of those changes on the results as well as on the accuracy of the data. The three major categories of changes that are discussed include 1) sampling and grouping; 2) IEP-log data collection and data entry; and 3) site visitation. In this section, we present the results of various exercises intended to show how differences in the methodologies may have influenced the estimates. It is important for the readers to note that the estimates in this section are exercises and should not to be interpreted as official results.

**Summary of Findings:** The IEP/service log findings across the two studies are best summarized in Table I-1, where it can be seen that the difference in agreement between the two years is substantial. For 2003-04, AIR found a 63.7% match between services as specified on the IEP and what was recorded on corresponding service logs across the special education population in LAUSD (excluding SLD), as compared to 93.2% from the PERB study. Likewise, the difference in AIR and PERB findings for students with SLD were also quite substantial.

**Table I-1. Percentages of Services Provided based on IEP-Log Data: Overall Population Estimates, 2003-04 (AIR) and 2004-05 (PERB and AIR)**

	Weighted to the Population with SLD A	Weighted to the Population without SLD B	Estimate for SLD Only C
2003-04 AIR	42.7%	63.7%	33.8%
2004-05 PERB	—	93.2%	72.8%
2004-05 AIR Analysis of PERB Data	84.9%	93.2%	72.8%

One interpretation of these data is there appears to have been substantial improvement in special education service delivery in the District across these two years as indicated by the substantial gains in IEP-log agreement shown above. However, we believe that several factors mitigate this conclusion. First, given the greater public awareness of the study and its methodology in this second year, we expected increases in IEP-log agreement from Year 1 due to better record-keeping and improved responsiveness to data requests. Improved responsiveness was also likely affected by the second factor: the repeated attempts to collect logs and the extension of the deadlines by which to submit them in Year 2, as compared to Year 1. Both factors would affect the comparability of these results as a true indicator of gains in the actual provision of service. However, it is possible that these results may be a better reflection of the services actually being provided due to providers responding to persistent requests and providing more complete data.

However, the direct involvement of the LAUSD Division of Special Education in obtaining logs in Year 2 may have created unintended pressure for providers to create logs where they were actually missing, or to show greater provision of service than actually occurred. Support for these concerns is found from the comparative analyses from the site visitation data, as shown in Table I-2.

**Table I-2. Percentages of Services Provided based on IEP-Site Visitation Data: Overall Population Estimates, 2003-04 (AIR) and 2004-05 (PERB and AIR)**

	Weighted to the Population with SLD	Weighted to the Population without SLD	Estimate for SLD Only
2003-04 AIR (Unmodified)	89.6%	82.6%	92.6%
2003-04 AIR (Modified*)	90.4%	85.6%	92.6%
2004-05 PERB	—	77.2%	79.0%
2004-05 AIR Analysis of PERB Data	79.2%	78.9%	79.4%

\* In the modified figures, provider absence due to jury duty, illness, or family emergency (n=11 in Year 1) count as evidence of service provision, according to an OIM decision regarding Year 2 data.

As shown above, the agreement between IEP and observed service provision as an indicator of actual special education service provision in LAUSD shows a fairly substantial drop between Year 1 and Year 2 based on these two studies (from 85.6% to 78.2%, excluding SLD). This is in direct contrast to the conclusion that might be drawn from Table 1, i.e., that service provision has substantially improved over these two years. However, as noted in the more detailed discussion at the end of this report, PERB methods in regard to the way in which the site visitations were conducted differed sufficiently from those employed by AIR to at least partially account for the differences shown above. For example, AIR returned for a second attempt to observe service provision if there were changes in the service schedule that prevented the service from being observed the first attempt. This approach would bias AIR results upward in relation to those reported by PERB, which coded those first attempts as non-provision of service.

Due to the relatively small size of the site visitation sample, the confidence intervals for these figures overlap and the differences noted above between the two years are not statistically significant. However, even though the results are not statistically significantly different, the site visitation data constitute an important check on the IEP-log analysis of service delivery, as with Year 1. As the IEP-log rate is significantly higher in Year 2 in comparison to the site visitation results, we interpret the IEP-log finding with some caution. At the same time, we recognize that the site visits are a more stringent measure of service delivery in that the service is to occur on a particular date and time. In this regard, we would expect to see a lower compliance rate in the site visitation than in the service log analysis, as the service may be provided at another time.

A final component of the IEP-log analysis is an assessment of agreement between the frequency and duration of service as shown on the logs and as specified in the IEPs. While differences are evident at the eligibility and service levels, the overall frequency rate did not change from Year 1 to Year 2, with both studies estimating agreement at 57.2% (unweighted). A similar trend was noted for duration, with overall estimates of 61.5% and 59.9% for Year 1 and Year 2, respectively. The fact that these overall rates are fairly similar across the two years, even with significant changes in the methodology, suggests that the rates are likely reasonable depictions of actual service delivery. The impact of the changes in the methodology is further explored in the report, with our conclusion that the Year 2 modifications did not adversely affect the reliability of the data in terms of frequency and duration.

**Summary Conclusions:** Based on the two years of analyses shown above, does it appear that the provision of special education service in LAUSD is improving or not? That is, are special education students in LAUSD receiving more special education services in accordance with what

is specified on their IEPs? In short, it appears that these findings are sufficiently ambiguous and inconsistent to say that there is a clear direction. There is evidence of improvement in recording services on provider logs and in having these logs available when repeatedly requested. To the extent that these studies create pressure for more accurate data reporting and collection, this improves the resulting estimates of the degree to which services are actually being provided. On the other hand, it should be recognized that too much pressure and repeated requests to show compliance in service provision may bias upwards the resulting estimates of services that are being received by special education students.

As described above, the IEP-log analysis over the two years suggests substantial improvement in service, while the IEP-site visit analysis does not substantiate such a conclusion. In both cases, it appears that changes in data collection methods may have affected the results as much, or perhaps greater than actual changes in the provision of service. We believe progress is being made in data collection and reporting and that the more IEPs and service logs are automated, the easier it will be to monitor progress across the full population of special education students in the District. However, at this point, we believe that continued site visits, as a cross check against what is shown in the service logs, will be needed.

**Recommendations:** We reviewed the five recommendations that PERB made in its Year 2 report and agree with each of these recommendations. Below we offer further suggestions to consider in conducting this study in Year 3.

Before embarking on Year 3, we recommend that PERB and AIR staff work together to establish common protocols based on what we have learned across both years and in which we have confidence. These agreed upon procedures will be used in subsequent years for comparability of results.

*Quality of Data:* We recommend that the service provider logs have a standard format (preferably electronic) that reports all relevant information needed to analyze service delivery in accordance to the IEPs (e.g., frequency and duration). We understand that LAUSD is moving towards an electronic tracking system for provider logs, which will likely provide the standardization required for a study of this magnitude. In establishing this system, it will be important to document reasons for provider absence, as well as the date that the information was entered and last revised. An electronic tracking system will allow for more real-time comparisons with the IEP, and should eliminate some of the problematic data collection and entry issues discussed in this report.

The substantial difference between what the logs show and what is observed in regard to service provision raises the possible question of log service validation. In many industries, service provision is cross-validated by obtaining a signature of the service recipient, as well as the notation of the service provider. In some cases, this might be the student, or perhaps the student's teacher in the case of younger children. Perhaps this is considered overly onerous for service providers by the district, but it would likely enhance the reliability of log data as a solid and consistently reliable indicator of special education service provision.

While the Welligent system has proved to be helpful in collecting IEP information, it is critical that the IEP has been properly and fully completed. Checks should be built into Welligent to ensure that the frequency and duration have been entered for each service. In regards to duration, additional checks should be implemented to ensure that the information recorded represents the total minutes required (instead of the minutes for each service session).

*Methodology:* Given the critical nature of this study, it is important that the data collection and entry process be as independent as possible. While the involvement of the Division of Special Education in collecting and cleaning the log data prior to it going to PERB may have alleviated some difficulties in obtaining logs and interpreting the information, it may raise questions about the results. In Year 3, we recommend that the logs go directly to PERB and that the involvement of the Division be minimal.

We believe that an important factor in affecting comparability of future data over time will be commonly adopted, and consistently maintained, conventions in regard to data collection. We recommend a *single request* for the logs with one follow-up and a limited window for submitting the logs. Providers should be well-aware by this time that they are required to maintain current logs documenting service provision. Repeated attempts to obtain logs may result in documents being manufactured after the fact, or in absence of service provision altogether. Similarly, with site visitation, the conditions under which the observer returns or does not return for a second chance to observe the provision of a missed service needs to be clearly and consistently specified and followed to allow comparable results across years.

In addition, the transparency of the data is crucial. The frequency and duration information needs to be entered into a database for analysis. We recognize that individual interpretation of the logs may in fact enhance the accuracy of the data, given the various ways of recording service information. However, we recommend that the frequency and duration information also be entered, as well as the data enterers' assessment of whether the log appeared in compliance with the IEP. If the analysis of the raw data conflict with the data enterers' assessment, these records would then be checked individually to resolve the discrepancy.

For tracking data enterer reliability, the person who first entered the IEP and log information should be identified in the database, as well as who checked the data and an indication if the information was changed. It is also important to establish a filing system for the IEP and log documents which will allow systematic validation activities to be conducted. In Year 1, we filed the documents by study ID, creating files which held both the IEP and logs for each student. Although logs with multiple students make this task difficult, we recommend that additional thought be given to the filing system for Year 3.

*Analysis:* The court may want to consider allowing a log with 1 missed session (in addition to valid reasons for non-provision of service) to still count as being compliant in the frequency analysis. For instance, if a service is required five times a week, this would amount to 40 service sessions over an 8-week period. If one session is not provided outside the allowed reasons,<sup>2</sup> the service for this student would be considered out of compliance. The table in Appendix C shows

---

<sup>2</sup> Reasons allowed for a service not provided are provider absence due to jury duty, illness, or family emergency; student absence; student no show; or school holiday.

the change in the frequency compliance rates if one deviation was allowed. This exercise, based on 244 log records, shows that the compliance rate would increase from 66.0% to 77.5%.

## II. Comparison of 2003-04 and 2004-05 Results and AIR’s Analysis of PERB Data

This section presents the 2003-04 results alongside two sets of 2004-05 results. The first set of estimates is taken directly from PERB’s Year 2 report, and the second set is derived from AIR’s analysis of PERB raw data. As described in the Year 2 report, PERB data enterers entered codes to indicate whether the log showed that the service provision complied or did not comply with the IEP requirements. Two other codes were used to denote records in which the data enterers could not determine whether the log was in compliance or to identify students who were dropped from the analysis. As shown in the following tables, we conducted our own analyses of these codes and compare them to the PERB estimates. Estimates by service category are shown in Appendix A.

### A. IEP-Log Analysis: Evidence of Service Provision:

Using logs as evidence, AIR estimated that 63.7% of the services in LAUSD were being provided in 2003-04 to students with all disabilities except Specific Learning Disabilities (SLD), which had an estimate of 33.8%. We acknowledged in the Year 1 report that the low estimate for SLD appeared to be more likely attributable to the poor quality of the Resource Specialist (RSP) logs than actual gaps in service. In Year 2, PERB shows substantial improvement in service provision, with 93.2% of the services being provided to the overall population (excluding SLD). Students with SLD also show an increased service rate of 73%. Our analyses of the PERB codes resulted in the same figures.

**Table II-1. Percentages of Services Provided based on IEP-Log Data: Overall Population Estimates, 2003-04 (AIR) and 2004-05 (PERB and AIR)**

	Weighted to the Population with SLD A	Weighted to the Population without SLD B	Estimate for SLD Only C
2003-04 AIR	42.7%	63.7%	33.8%
2004-05 PERB	—	93.2%	72.8%
2004-05 AIR Analysis of PERB Data	84.9%	93.2%	72.8%



**Table II-2. Evidence of Service Provision, 2003-04 (AIR) and 2004-05 (PERB and AIR), by Disability**

Disability	2003-04 AIR		2004-05 PERB		2004-05 AIR Analysis of PERB Data <sup>1</sup>	
	N	%	N	%	N	%
Autism	350	68.1	501	94.9	501	94.9
Deaf/Hard of Hearing	367	77.4	518	94.9	518	94.9
Emotional Disturbance	121	39.3	248	81.1	248	81.0
Mental Retardation	414	69.6	369	95.8	369	95.8
Multiple Disability/Deaf-Blind	279	75.4	424	98.2	424	98.1
Orthopedic Impairment/ Traumatic Brain Injury	213	60.6	643	92.8	643	92.8
Other Health Impairment	424	55.0	348	83.7	348	83.7
Specific Learning Disability	78	33.8	526	72.8	526	72.8
Speech and Language Impairment	206	65.8	321	95.0	321	95.0
Visual Impairment	289	65.8	615	97.6	615	97.6
<b>Total<sup>2</sup></b>	<b>2,741</b>	<b>63.5</b>	<b>4,513</b>	<b>90.3</b>	<b>4,513</b>	<b>90.3</b>
<b>Total Service Records Analyzed</b>	<b>4,316</b>		<b>4,997</b>		<b>4,997</b>	

<sup>1</sup> AIR conducted analyses on PERB data.

<sup>2</sup> Total figures include SLD. In the Year 2 report for 2004-05, PERB did not include SLD in the total (3,987 records, 93.3%).

## B. IEP-Log Analysis: Frequency Compliance

A second component of the study was to assess whether services were being provided at the frequency required by the IEPs. While we did not provide a population estimate for frequency compliance in 2003-04, Table II-3 presents this information for the population excluding SLD (52.8%) and SLD only (51.5%) for 2004-05, using AIR’s analysis of PERB data.

**Table II-3. Percentages of Services Meeting IEP Frequency, Overall Population, 2004-05 (AIR)**

	Weighted to the Population with SLD	Weighted to the Population without SLD	Estimate for SLD Only
2004-05 AIR Analysis of PERB Data	52.2%	52.8%	51.5%

Table II-4 shows that there is no change in the overall (unweighted) frequency compliance rate from Year 1 to Year 2 (57.2%), although students with Emotional Disturbance and Speech/Language Impairment showed declines of 20 and 12 percentage points.

The reader will note that AIR’s analysis of the PERB data in Table II-4 (last column) produces slightly different numbers. In the overall total number of records analyzed, we included records in which 1) PERB “could not determine whether frequency was met”<sup>3</sup> and 2) there was frequency information in the IEP, thereby lowering the compliance percentages.

<sup>3</sup> PERB used this code when either the IEP lacked frequency or when the data enterer could not determine if the information in the log met the IEP requirements. PERB did not include these records in the overall number of service records analyzed.

**Table II-4. Percentages of Services Meeting IEP Frequency, 2003-04 (AIR) and 2004-05 (PERB and AIR), by Disability**

Disability	2003-04 AIR		2004-05 PERB		2004-05 AIR Analysis of PERB Data <sup>1</sup>	
	N	%	N	%	N	%
Autism	174	58.2	258	56.3	256	56.3
Deaf/Hard of Hearing	154	55.6	281	57.8	281	57.9
Emotional Disturbance	64	70.3	101	49.0	101	48.6
Mental Retardation	137	56.6	183	53.6	182	52.4
Multiple Disability/Deaf-Blind	182	54.0	218	60.1	218	59.4
Orthopedic Impairment/ Traumatic Brain Injury	187	59.0	363	66.9	360	66.5
Other Health Impairment	89	53.3	172	56.4	171	55.7
Specific Learning Disability	23	51.1	231	52.3	227	51.5
Speech and Language Impairment	106	60.9	141	48.8	140	48.3
Visual Impairment	142	56.6	345	60.4	345	60.0
<b>Total<sup>2</sup></b>	<b>1,258</b>	<b>57.2</b>	<b>2,293</b>	<b>57.2</b>	<b>2,281</b>	<b>56.8</b>
<b>Total Service Records Analyzed</b>	<b>2,200</b>		<b>4,011</b>		<b>4,016</b>	

<sup>1</sup> AIR conducted analyses on PERB data.

<sup>2</sup> Total figures include SLD. In the Year 2 report for 2004-05, PERB did not include SLD in the total (2,062 records, 57.8%).

### C. IEP-Log Analysis: Duration Compliance

The Year 1 and 2 studies also determined the rates at which services were provided at the total monthly duration required by the IEP. As with the frequency analysis, AIR did not provide a population estimate for this component in Year 1. Table II-5 shows that 55.6% of the services were provided to the overall population (excluding SLD) in 2004-05, using the data collected by PERB.

With respect to the unweighted percentages, there was only a marginal decline from 2003-04 (61.5%) to 2004-05 (59.9%). Although there were differences in the methodology (discussed further in Section III), these figures suggest that no progress has been made in terms of duration compliance based on the log data. Again, AIR’s analysis of the PERB data generate slightly different estimates due to the fact that records with IEP duration but for which PERB could not determine compliance were included in the total number of records analyzed.

**Table II-5. Percentages of Services Meeting Total Duration, Overall Population, 2004-05 (AIR)**

	Weighted to the Population with SLD	Weighted to the Population without SLD	Estimate for SLD Only
2004-05 AIR Analysis of PERB Data	54.9%	55.6%	54.1%

**Table II-6. Percentages of Services Meeting Total Duration, 2003-04 (AIR) and 2004-05 (PERB and AIR), by Disability**

Disability	2003-04 AIR		2004-05 PERB		2004-05 AIR Analysis of PERB Data <sup>1</sup>	
	N	%	N	%	N	%
Autism	186	64.6	266	58.1	265	57.9
Deaf/Hard of Hearing	108	63.5	290	59.9	290	59.9
Emotional Disturbance	68	77.3	130	65.0	130	64.0
Mental Retardation	137	59.8	190	55.1	189	54.9
Multiple Disability/Deaf-Blind	171	58.2	217	60.4	217	59.5
Orthopedic Impairment/ Traumatic Brain Injury	172	60.8	366	67.5	364	67.3
Other Health Impairment	96	60.0	177	59.2	176	57.7
Specific Learning Disability	23	53.5	242	55.6	240	54.1
Speech and Language Impairment	96	56.1	146	50.7	145	50.0
Visual Impairment	96	64.0	358	63.1	358	62.5
<b>Total<sup>2</sup></b>	<b>1,153</b>	<b>61.5</b>	<b>2,382</b>	<b>59.9</b>	<b>2,374</b>	<b>59.2</b>
<b>Total Service Records Analyzed</b>	<b>1,876</b>		<b>3,977</b>		<b>4,007</b>	

<sup>1</sup> AIR conducted analyses on PERB data.

<sup>2</sup> Total figures include SLD. In the Year 2 report for 2004-05, PERB did not include SLD in the total (2,140 records, 60.4%).

#### D. Service Provision: Site Visitation

Both studies incorporated a site visitation component to serve as a check on the estimates generated by the IEP-log analysis. Although these are not intended to be generalized to the overall population, Table II-7 shows that the population estimates (excluding SLD) declined slightly from 85.6% to 77.2%. However, the confidence intervals for these figures overlap, suggesting no significant differences.

**Table II-7. Percentages of Services Provided based on IEP-Site Visitation Data: Overall Population Estimates, 2003-04 (AIR) and 2004-05 (PERB and AIR)**

	Weighted to the Population with SLD	Weighted to the Population without SLD	Estimate for SLD Only
2003-04 AIR (Unmodified)	89.6%	82.6%	92.6%
2003-04 AIR (Modified*)	90.4%	85.6%	92.6%
2004-05 PERB	—	77.2%	79.0%
2004-05 AIR Analysis of PERB Data	79.2%	78.9%	79.4%

\* In the modified figures, provider absence due to jury duty, illness, or family emergency (n=11 in Year 1) count as evidence of service provision, according to an OIM decision regarding Year 2 data.

Please note that the Year 1 figures have been modified to reflect a decision made by the Office of the Independent Monitor (OIM) regarding the interpretation of certain provider absences in Year 2. If the provider was absent due to jury duty, illness, or family emergency, the service was considered to have been provided. The 11 observations for which this occurred in Year 1 are treated as evidence of service provision in Table II-8 below. Even *without* this modification, the percentage of services provided in Year 1 is higher than what was observed in Year 2. As with the population estimates, the confidence intervals for the unweighted figures overlap.

**Table II-8. Number of Percentage of Site Visitation Observations by Session Status, 2003-04 (AIR) and 2004-05 (PERB)**

Status of Session	AIR Observations 2003-04 MODIFIED	AIR % 2003-04 MODIFIED	PERB Observations 2004-05 <sup>&lt;1&gt;</sup>	PERB % 2004-05 <sup>&lt;1&gt;</sup>
<b>Service Provided</b>	<b>288</b>	<b>84.7%</b>	<b>242</b>	<b>74.5%</b>
Code 1. Session Completed	226	66.5%	185	56.9%
Code 2. Service Provided but Session Incomplete	18	5.3%	23	7.1%
Code 3. Provider Absent (illness, emergency, jury duty) <sup>&lt;2&gt;</sup> MODIFIED	11	3.2%	9	2.8%
Code 5. Student Absent/Provider Present	27	7.9%	22	6.8%
Code 6. Student No Show/ Provider Present	6	1.8%	3	0.9%
<b>Service Not Provided<sup>&lt;2&gt;</sup></b>	<b>52</b>	<b>15.3%</b>	<b>83</b>	<b>25.5%</b>
Code 4. Provider in Meeting/ Student Present	17	5.0%	10	3.1%
PERB Code 7. Provider Absent (reason unknown)/ Student Present	n/a	n/a	41	12.6%
AIR Code 7 / PERB Code 8. Service Not provided <sup>&lt;3&gt;</sup>	35	10.3%	32	9.8%
<b>Total</b>	<b>340</b>		<b>325</b>	

<sup>1</sup> AIR conducted an analysis on the PERB coded data and generated the same results.

<sup>2</sup> The original figures derived from Year 1 have been modified in this table, per the clarification provided by OIM following the Year 1 results. In the original report, the 11 observations in which the provider was absent due to illness, family emergency, or jury duty were treated as "service not provided" (resulting in a service provision rate of 81.5%). In this table, these observations have been re-categorized as valid evidence of service provision.

<sup>3</sup> This category in Year 1 included "other" situations, such as provider no show or no provider on staff. See Table 3.17 in the Year 1 report.

## **I. Changes from Year 1 to Year 2**

This section reviews the changes in the methodologies from Year 1 to Year 2 and their possible impacts on the results shown in Section II. The changes are grouped under three major categories: sampling and grouping issues; IEP-log data collection and data entry issues; and site visitation. The tables below briefly outline the differences in the Year 1 and Year 2 methodologies, effect on compliance rates, the expected direction of the possible effect (if a direction can be estimated), and the estimated effect that the change could have had on the accuracy of the data. Specific issues which require additional discussion, or for which we conducted exercises, are explained further in each sub-section (and are boldfaced in the issue tables).

**A. Sampling and Grouping Issues**

<b>SAMPLING</b>					
Issue	Differences between Year 1 and Year 2		Effect on Compliance Rate	Direction of Effect	Estimated Effect on Accuracy of Data
	Year 1	Year 2			
1. Mapping of Eligibility	MDV collapsed with MD	MDV collapsed with VI (117 MDV students affected)	Minor	Negligible. Weighted % of services provided (excluding SLD) increased marginally from 93.1% to 93.2%	No change
	DD and DI collapsed with OHI	DD and DI collapsed with MR (86 DD and 12 DI students affected)	Minor		
2. Mapping of Services	NPA Speech collapsed with Language & Speech	NPA Speech (n=81) collapsed with NPA Services	Minor	Negligible. Weighted % of services provided (excluding SLD) decreased marginally from 93.1% to 93.0%	No change
	Pre-Kindergarten Itinerant (PKIT) grouped with Language/ Speech	PERB created a new service category to capture preschool services (PKIT, ASSEP, PHONO, ILEAP, PKIT-HS). These observations (n=95) were subsequently dropped from the site visit component.			
3. Sample Size	1 <sup>st</sup> stage: 3,800, 380 in each disability category Target: 3,300 Actual: 2,997	1 <sup>st</sup> stage: 4,500, 380 in each disability category and 1,080 for SLD Target: 3,800 Actual: 4,134	Likely/Minor	Unknown	More accurate. The large sample size of the first stage sample provides an accurate estimate under both approaches. The larger sample size for SLD improves the precision of the estimate for this group of students.
	2 <sup>nd</sup> stage: 410, 40 in each disability category, except OTH/TBI (50) Target: 330 Actual: 270	2 <sup>nd</sup> stage: 460, 40 in each service category except APE and RSP (50) Actual: 325	Likely/Minor	Unknown	No change.
4. Site Visitation Sampling	<b>Quota sample: 40 by disability category, except for OTH/TBI with 50. Eligibility was main stratifier.</b>	<b>Not a quota system by eligibility. Still maintained purpose of obtaining 30 observations, but priority given to services rather than eligibility.</b>	Likely/Minor	<b>Unknown. The small sample size generates large confidence intervals. Therefore, small changes in the sampling strategy are not likely to generate statistically significantly different results than the ones obtained in Year 1.</b>	No change.

**Sampling Issue 4: Site Visitation Sampling**

The sampling design used to select students for the site visitation observations in Year 2 was not entirely clear. Although the Year 2 design was intended to obtain 30 observations by disability (i.e., the main stratifier in Year 1) and by service category, priority was given to the latter. Each service had 40 records in the sample, except for Adaptive PE (APE) and Resource Specialist (RSP) with 50 records.<sup>4</sup> However, many students fall out of the sample for various reasons (e.g., service no longer provided per new IEP, left the district). Given that no over-sampling was built into the Other Health Impairment (n=30) category, it would have been very difficult to have met the goal of 30 observations for this disability category. Indeed, as shown in Table III-1, observations were completed for only 19 students with OHI. Although the actual numbers of obtained observations by disability category were generally higher in Year 2, we recommend that the focus of the sampling be on the disability category in Year 3, including a certain over-sampling rate.

**Table III-1. Sample and Number and Percentages of Site Visit Observations, by Disability, 2003-04 (AIR) and 2004-05 (PERB)**

Disability	AIR Sample 2003-04	PERB Sample 2004-05	AIR Observations (students counted once) 2003-04		PERB Observations 2004-05	
	N	N	N	%	N	%
AUT	40	46	29	10.7	35	10.8
DHH	40	44	26	9.6	38	11.7
ED	40	41	19	7.0	30	9.2
MD/DB	40	46	32	11.9	35	10.8
MR	40	36	27	10.0	25	7.7
OHI	40	30	23	8.5	19	5.8
OI/TBI	50	61	32	11.9	36	11.1
SLD	40	42	26	9.6	34	10.5
SLI	40	56	28	10.4	26	8.0
VI	40	58	28	10.4	47	14.5
<b>Total</b>	<b>410</b>	<b>460</b>	<b>270</b>	<b>100</b>	<b>325</b>	<b>100</b>

<sup>4</sup> In Year 1, 40 students were sampled from each disability category, except for Orthopedic Impairment/Traumatic Brain Injury. Fifty students were drawn from each of these two categories due to the low incidence of services received by these students.

**B. IEP-Log Data Collection and Data Entry Issues**

<b>IEP-LOG DATA COLLECTION &amp; DATA ENTRY</b>					
<b>Issue</b>	<b>Differences between Year 1 and Year 2</b>		<b>Effect on Compliance Rate</b>	<b>Direction of Effect</b>	<b>Estimated Effect on Accuracy of Data</b>
	<b>Year 1</b>	<b>Year 2</b>			
1. Log requests – Different basis	Initial requests based on SESAC; follow-up request to capture services in IEP but not SESAC	Requests were based entirely upon IEP information	Likely/Minor	Possibly increase rates	More accurate
2. Log requests – Multiple requests and follow-ups	Requests made at two points – Fall 2003 (SESAC) and January 2004 (IEP)	Requests made twice, to follow up on missing logs	Likely/ Possibly Major	Possibly increase in rates	Mixed
3. Log requests – Division of Spec Ed	Requests made through the OIM	Division of Spec Ed involved in the data collection	Likely/ Possibly Major	Possibly increase in rates	Mixed
4. Log requests - Time-Sensitive/8-week period	Request made for 1 month only	Requests made for 2 months, specific 8-week period	Likely/ Significant	Using 1-month of logs for a sample of students, frequency compliance rate <b>declined</b> from 59.9% to 41.9% (See Table III-2)	Mixed
5. Use of Welligent IEP database	Paper copies of all IEPs were collected and entered by hand	For 54% of sample, extract from the Welligent database for online IEPs were used. Information for students with batch-entered Welligent and paper IEPs were entered by hand.	Likely/Possibly Minor	Possibly increase rates, if schools participating in Welligent have better service delivery than schools not in Welligent.	More accurate IEP info
6. Frequency and duration not entered	Data enterers entered frequency and duration information on log; information analyzed with Stata program	Log information on frequency/duration not entered. Data enterers judged whether service is in compliance with IEP	Highly Likely/ Possibly Major	Analysis of 269 log records shows that PERB interpretation produces a <b>slightly higher</b> compliance rate. See Tables III-3 to III-6	Mixed
7. Duration interpretation	Duration listed in IEP was considered “total” – not the duration of individual sessions.	Data enterers told that IEP weekly duration could be session minutes or weekly minutes.	Highly Likely/ Possibly Major for Certain Services	Possibly decrease rates. See Tables III-7 and III-8.	Mixed

**Data Collection Issue 3: Log requests – Involvement of the Division of Special Education**

In Year 1, all data requests were made through the OIM. In Year 2, PERB directed its follow-up requests for logs to the Division of Special Education. While the Division may have been better able to obtain providers’ logs than other entities, concerns are raised about unintended pressure this may have created for providers to manufacture logs after the fact, thereby increasing rates of documented service delivery. While one might contend that the timing of completing the logs themselves is secondary to whether the service was actually provided, log records need to be



current and up-to-date in order for any accountability system to work effectively. In addition, we were informed that the Division “cleaned” the log information before providing it to PERB to make it easier to interpret. While this ultimately may have had no impact on the estimates, it is important that the data collection and data entry process remain as objective as possible, particularly in a high-stakes study such as this.

#### **Data Collection Issue 4: Log requests -Time-Sensitive/8-Week Period**

In Year 1, we requested a single month of logs, whereas in Year 2, providers were to provide logs for a specific 8-week period. We expected this larger window to result in a decline in compliance rates. For instance, if the logs provided did not cover the exact 8-week time period requested, the service was automatically out of compliance for frequency and duration. Furthermore, the bar is higher: if one session was not provided across 8 weeks, except for permitted reasons, service is considered out of compliance under the Year 2 approach.

To assess the impact of this change on frequency and duration rates, we entered and analyzed data from 269 randomly selected service logs received by PERB. The data were entered according to the specific 8-weeks requested by PERB and analyzed for the 8-week period.<sup>5</sup> For the single month analysis, we analyzed these data from the October or January logs, as they represented the fullest months within their respective 8-week timeline.

Contrary to our expectations, the frequency compliance rate for the 8-week period was higher overall than the rate derived from a single-month analysis, and particularly so for APE and VI services (Table III-2). With respect to APE, the low compliance rate from the single-month analysis stems from the fact that January data entered for our sample did not reflect holidays occurring at the beginning of the month. APE is often a daily service, and hence would be highly impacted by this missing information.

The results of this exercise allay concerns regarding the potential negative impact of using an 8-week period to assess frequency compliance rate. It is also possible that providers are more able to provide compensatory sessions during this larger window to make up for missed sessions; these sessions may not necessarily be captured in a single month timeframe.

---

<sup>5</sup> Of the 269 logs entered, 25 records did not have a frequency in the IEP. This resulted in a sample of 244 records for this exercise.

**Table III-2. Comparison of Frequency Compliance Rates Using Information from a Single Month (AIR data) versus Specific 8-Week Period (AIR data)**

	Frequency Compliance Rate with 1 Month	Frequency Compliance Rate with 8-Weeks
	%	%
<b>APE</b>	<b>15.9</b>	<b>64.4</b>
AUD	66.7	58.8
<b>VI</b>	<b>9.1</b>	<b>50.0</b>
D/HH	55.6	77.8
LRE	72.7	85.2
LAS	23.8	45.5
SMH	33.3	53.3
NPA	75.0	73.3
OT	47.1	65.7
PKIT	85.7	85.7
PT	58.3	76.9
RSP	50.0	59.1
<b>Total</b>	<b>46.0</b>	<b>66.0</b>
<b>Total Service Records Analyzed</b>	<b>237*</b>	<b>244</b>

\* Total number of service records lower for 1 month analysis, as some records may be missing October or January logs

#### **Data Entry Issue 6: Frequency and Duration Not Entered**

When producing estimates for a high-stakes purpose such as this, transparency of data is critical for ensuring the reliability of the analyses through replication. In Year 1, AIR data enterers entered the frequency and duration information from the logs into a database, so the data analyzed were actually a reflection of the information contained in the logs. This was later analyzed against the frequency and duration required by the IEP. In a departure from this approach, PERB data enterers were trained to interpret during their data entry whether the log information was in compliance with the IEP. For each component (evidence of service, frequency, and duration), the data enterers would enter a “1” to indicate that the log met the IEP requirements, a “2” when the log did not meet the IEP requirements, and a “3” when it could not be determined. No frequency or duration information was entered. Consequently, under this approach it is only possible to analyze the *PERB interpretations* for all of the logs entered (e.g., tabulations of code 1’s, etc. - see Section II), but not to replicate the actual analysis using raw data of the logs. This approach raises questions as to how accurately the data enterers interpreted the log information, and as to how it is possible (whether at all) to replicate the results.

Although we were unable to conduct independent frequency and duration analyses on all logs, we entered information for 269 randomly selected logs and assessed the extent to which our estimates matched with the PERB codes. Table III-3 shows that AIR frequency analysis matched with PERB’s codes in 89% of the cases.

**Table III-3. Number and percentage of sampled service records for which AIR and PERB analysis showed agreement in frequency analysis (n=269)**

	N	%
AIR frequency analysis did <b>not</b> agree with PERB	30	<b>11.2</b>
AIR frequency analysis did agree with PERB	239	<b>88.9</b>
<b>Total</b>	<b>269</b>	<b>100</b>

\* Numbers do not add to 100 percent due to rounding.

Table III-4 disaggregates these figures further by showing exactly where the discrepancies appeared. The boldface figures running diagonally are the 239 records in which the AIR and PERB numbers agreed (239 = 152 + 64 + 22). For a majority of the 30 records with discrepancies (19 of 30), AIR determined that the frequency did *not* meet the IEP requirement whereas PERB indicated that the record did meet compliance. Overall, PERB found that 63.6% of the 269 records complied with the IEP (yellow cells), whereas AIR estimated that 59.9% of the records were in compliance (green cells).

**Table III-4. Number and percentage of sampled service records for which AIR and PERB analysis showed agreement in frequency analysis, by compliance category (n=269)**

		AIR Frequency Analysis			PERB	
		Met Compliance	Did Not Meet Compliance	Can't Determine	Total N	PERB %
<b>PERB Frequency Analysis</b>	Met Compliance	<b>152</b>	19	0	<b>171</b>	<b>63.6</b>
	Did Not Meet Compliance	8	<b>64</b>	3	<b>75</b>	<b>27.9</b>
	Can't Determine	1	0	<b>22</b>	<b>23</b>	<b>8.6</b>
<b>AIR Total N</b>		<b>161</b>	<b>83</b>	<b>25</b>	<b>269</b>	<b>100</b>
<b>AIR %</b>		<b>59.9</b>	<b>30.9</b>	<b>9.3</b>		<b>100</b>

The same exercise was completed for the duration component. Table III-5 shows a higher discrepancy rate of 14.6% for duration in relation to the frequency comparisons above. For the 268 records analyzed, AIR and PERB agreed in 85.5% of the cases.<sup>6</sup>

<sup>6</sup> If the total service time provided over the 8 weeks was 10 minutes or less below the total amount of time required by the IEP for an 8-week period, AIR counted this as being in compliance.

**Table III-5. Number and percentage of sampled service records for which AIR and PERB analysis showed agreement in duration analysis (n=268)**

	N	%
AIR duration analysis did <b>not</b> agree with PERB	39	<b>14.6</b>
AIR duration analysis did agree with PERB	229	<b>85.5</b>
<b>Total</b>	<b>268</b>	<b>100*</b>

\* Numbers do not add to 100 percent due to rounding.

As with the above frequency table, Table III-6 shows where the AIR analysis and PERB interpretation of the duration agreed and disagreed. Again, the diagonal boldface numbers indicate the matches between AIR and PERB (229 = 163 + 42 + 24). Among the 39 records with contrary interpretations, AIR found 21 records did not meet compliance, while PERB indicated that they did meet the IEP requirement. In 15 cases, the opposite occurred, with PERB coding the record as not meeting compliance and the AIR analysis indicated the log was in compliance. Overall, PERB found that 68.7% of the sampled logs demonstrated compliance with the IEP information, with AIR showing 66.8% compliance rate for the same logs.

**Table III-6. Number and percentage of sampled service records for which AIR and PERB analysis showed agreement in duration analysis, by compliance category (n=268)**

		AIR Duration Analysis			PERB Total N	PERB %
		Met Compliance	Did Not Meet Compliance	Can't Determine		
<b>PERB Duration Analysis</b>	Met Compliance	<b>163</b>	21	0	<b>184</b>	<b>68.7</b>
	Did Not Meet Compliance	15	<b>42</b>	2	<b>59</b>	<b>22.0</b>
	Can't Determine	1	0	<b>24</b>	<b>25</b>	<b>9.3</b>
<b>AIR Total N</b>		<b>179</b>	<b>63</b>	<b>26</b>	<b>268</b>	<b>100</b>
<b>AIR %</b>		<b>66.8</b>	<b>23.5</b>	<b>9.7</b>		<b>100</b>

The high degree of alignment between AIR’s analysis and PERB’s codes for frequency and duration for this limited sample are encouraging. Based on these comparisons, PERB’s interpretations appear to have been reasonable and consistent for a majority of the records. Without a standardized format, we recognize that logs are generally “messy” and do not easily conform to data entry, and that the variable nature of the logs may require some individual interpretation to accurately assess compliance. However, given the purpose of the study, we strongly recommend that the data be as transparent as possible. One alternative to consider for Year 3 is to enter the frequency and duration information, as well as enter a compliance code based on the data enterers’ assessment. Comparisons can then be made between estimates derived from a program analyzing the IEP-log data and the data enterers’ codes. Discrepancies between the two (program vs. data enterer codes) can be followed up and resolved individually as necessary.

### Data Entry Issue 7: Duration Interpretation

Using the log information as a guide, PERB data enterers were advised to use their judgment as to whether the weekly duration listed in the IEP represented the required total weekly duration, or reflected the duration required for each individual session. That is, if the IEP lists a frequency of 2 times a week and a duration of 30 minutes per week, the data enterer could interpret the required weekly duration as 30 minutes (as listed), or 60 minutes (2 x 30). This could result in substantially different requirements over an 8-week period (240 versus 480 minutes). In Year 1, we did not allow for this variation. The weekly duration cited in the IEP was taken to be the total required weekly duration (or 30 minutes, in the above example).

Tables III-7 and III-8 below show the services and disability categories with weekly frequencies greater than 1 in the IEP, based on the sample of IEPs drawn in Year 2. APE and RSP are the most impacted services, with the vast majority of these services records showing a weekly frequency greater than 1. Among the 765 APE records with a frequency greater than 1, 43% had a frequency of 2 and a duration of 60 minutes, leaving the interpretation of whether the required weekly duration was 60 or 120 minutes to the data enterer. Over a quarter (28%) of the 765 APE records had a frequency of 5 and a weekly duration of 250 minutes, which is less subject to interpretation (as it is impractical for a single APE session to last 250 minutes). The RSP records with a weekly frequency greater than 1 showed less clustering. Only 17% of these had a frequency of 5 and a weekly duration of 250 minutes; all other records had various frequencies and durations. This suggests that the total required duration for RSP may have been more subject to varying interpretations.

**Table III-7. Service records with weekly frequency greater than 1 (excluding records in which frequency or duration was missing or zero), by service\***

	Frequency Equals 1		Frequency Greater Than 1		Total	
	N	%	N	%	N	%
<b>APE</b>	235	23.5	<b>765</b>	<b>76.5</b>	1,000	100
DHH	326	95.9	14	4.1	340	100
LAS	889	83.9	171	16.1	1,060	100
LRE	14	73.7	5	26.3	19	100
NPA	47	68.1	22	31.9	69	100
OT	369	98.9	4	1.1	373	100
PRE	10	83.3	2	16.7	12	100
PT	35	97.2	1	2.8	36	100
<b>RSP</b>	151	19.9	<b>607</b>	<b>80.1</b>	758	100
SMH	363	96.5	13	3.5	376	100
VI	197	59.9	132	40.1	329	100
Total	2,636	60.3	1,736	39.7	4,372	100

\* Information was taken from the sample of IEPs drawn for this study and does not reflect service records for which logs were provided (e.g., those that were analyzed)

With respect to disability category, students with Orthopedic Impairment/Traumatic Brain Injury, Specific Learning Disability, and Visual Impairment showed high percentages of services with a weekly frequency greater than 1.

**Table III-8. Service records with weekly frequency greater than 1 (excluding records in which frequency or duration was missing or zero), by disability\***

	Frequency Equals 1		Frequency Greater Than 1		Total	
	N	%	N	%	N	%
AUT	386	69.2	172	30.8	558	100
DHH	381	79.4	99	20.6	480	100
ED	224	83.3	45	16.7	269	100
MD/Deaf-Blind	191	57.0	144	43.0	335	100
MR	224	61.5	140	38.5	364	100
OHI	189	53.2	166	46.8	355	100
<b>OT/TBI</b>	218	47.5	<b>241</b>	<b>52.5</b>	459	100
<b>SLD</b>	238	39.1	<b>370</b>	<b>60.9</b>	608	100
SLI	271	87.4	39	12.6	310	100
<b>VI</b>	314	49.5	<b>320</b>	<b>50.5</b>	634	100
<b>Total</b>	2,636	60.3	1,736	39.7	4,372	100

\* Information was taken from the sample of IEPs drawn for this study and does not reflect service records for which logs were provided (e.g., those that were analyzed)

There are two possible effects of this change in interpreting the IEP duration information. If the IEP was completed incorrectly (i.e., the weekly duration *should* be multiplied by the frequency in order to calculate the real required duration), this individual interpretation could improve the accuracy of the analysis. On the other hand, if the IEP correctly notes the required duration, multiplying the duration by the frequency could erroneously penalize the district, as the expected level of service would be higher than what is actually required. Given that this issue appears to be isolated to particular services and disabilities, we do not believe that this change in methodology has large implications for the population estimates. However, we recommend a consistent approach to avoid potential variations in how data enterers should interpret the data. Ideally, IEP clerks would be appropriately trained to enter the information correctly and that a check be built into the Welligent system to ensure that the duration entered represents the total amount of time that the service is to be provided.

**C. Site Visitation Issues**

<b>SITE VISITATION</b>					
<b>Issue</b>	<b>Differences between Year 1 and Year 2</b>		<b>Effect on Compliance Rate</b>	<b>Direction of Effect</b>	<b>Estimated Effect on Accuracy of Data</b>
	<b>Year 1</b>	<b>Year 2</b>			
1. Flexible Services	If the designated service was too flexible, we attempted to keep the student in the sample by observing a different service	Students whose designated service was too flexible were dropped from the sample entirely	Likely/ Possibly Minor	Possibly increase rates. Students with flexible services may have different service delivery rates than students with fixed schedules.	Unknown
2. Inclusion of NPA Services	NPA services were not included in Year 1 site visits	NPA services were included in Year 2 site visits	Likely/Minor	Possibly increase rates. In Year 2, nearly 95% of students with NPA services had evidence of service provision, and more than 80% of NPA logs showed duration compliance (in the IEP-log analysis).	More accurate
3. Use of School Site IEP	If school reported that a newer IEP stated the student was no longer receiving the service, we attempted to observe another service received by that student	If the more recent school-site IEP indicated that the student did not receive the service, the student was dropped from the sample	Likely/Minor	Unknown	More accurate
4. Different use of codes	<b>Code 3 (jury duty, illness, emergency) was treated as service not provided</b>	<b>Code 3 was treated as legitimate service provision.</b>	<b>Likely/Minor</b>	<b>Increase rates. When re-categorizing Year 1 results so that Code 3 is considered valid service provision, service provision rate increased from 81.5% to 84.7%.</b>	<b>More accurate, based on court decision</b>
	<b>Seven codes</b>	<b>Eight codes</b>	<b>Possibly Minor</b>	<b>Unknown</b>	<b>Unknown</b>
	<b>If service schedule changed from what was provided over the phone, site visitors attempted another observation</b>	<b>If service schedule changed from what was provided over the phone, the observation was coded as 8 (service not provided).</b>	<b>Likely/ Possibly Minor</b>	<b>Decrease rates</b>	<b>Less accurate</b>
5. Duplicate observations	To increase the number of observations for low incidence services, site visitors observed a student more than once	No student was observed more than once	Minor	If we used the first observation for each student (e.g., students counted once only) in Year 1, the percentages of students with service provision would have <u>increased slightly</u> from 84.7% to 85.6% (See Appendix B).	Mixed

#### **Site Visitation Issue 4: Different Use of Codes**

In Year 2, three changes were made to the use of codes for the site visitation component. Per a decision made by OIM, provider absences due to jury duty, illness, or family emergency would be treated as valid evidence of service delivery in Year 2. To assess the impact of this change, we re-categorized Year 1 observations coded in this manner. As shown in Section II, the service delivery rate increased from 81.5% to 84.7%, which increased the gap between Year 1 and Year 2 (74.5%) results.

Another modification was made to the coding system used in Year 2, which used eight codes in comparison to seven in Year 1. The meaning of the codes 1-6 align between the two years. Code 7 in Year 1 reflected situations in which no service was provided, including provider absence not related to jury duty, illness, or family emergency. PERB redefined code 7 to denote situations in which the provider was absent (reason unknown) and the student was present. Code 8 was then added to reflect other situations in which service was not provided. In and of itself, this change should not have affected the results, as these codes represent observations in which service was not provided. One would expect the percentages of code 7 in Year 1 (10.3%) to be somewhat similar to the combined percentage of codes 7 and 8 in Year 2 (22.4%), but this is not the case.

Upon closer examination of the observation notes for code 7's in Year 2, we found seven instances in which we would have followed up with a second attempt to observe according to our Year 1 approach (i.e., if the visitor was told that the service already took place or would take place at a different time/day, the visitor was to return for a second attempt). It also appears that being persistent in obtaining a set time for normally flexible services in Year 2 may have contributed to the higher rates of services coded as 7 (i.e., services not provided). In Year 1, if the designated service was too flexible to observe (flexible day and flexible time), AIR replaced the service with another service received by the same student. If that was not an option, the student was then replaced. A review of the observation notes for code 7 indicates that visitors attempted to observe several flexibly provided services. PERB noted in its report that several attempts were made to obtain a day and time for flexible services, if the service was not generally provided at that time.

While the overall percentage of service delivery is lower in Year 2 in comparison to Year 1, these observation and coding strategies may account for some of the difference. As mentioned, the confidence intervals overlap for the two estimates, indicating that the difference is not statistically significant.



**Appendix A**

**Table 1A. Evidence of Service Provision, 2003-04 (AIR) and 2004-05 (PERB and AIR), by Service**

Service	2003-04 AIR		2004-05 PERB		2004-05 AIR Analysis of PERB Data <sup>1</sup>	
	N	%	N	%	N	%
Adaptive PE	726	68.9	965	98.8	965	98.8
Deaf/Hard of Hearing	283	82.7	389	99.7	389	99.7
Language and Speech	781	75.5	1,030	95.8	1,030	95.8
LRE	107	50.0	168	85.3	168	85.3
Mental Health	136	41.0	358	87.5	358	87.5
Non-Public Agency	22	55.0	52	94.5	52	94.5
Occupational Therapy	294	77.0	395	98.3	395	98.3
Physical Therapy	85	62.0	131	100.0	131	100.0
Pre-School	n/a	n/a	75	100.0	75	100.0
RSP	120	22.6	625	65.2	625	65.2
Visual Impairment	187	74.2	325	99.4	325	99.4
<b>Total</b>	<b>2,741</b>	<b>63.5</b>	<b>4,513</b>	<b>90.3</b>	<b>4,513</b>	<b>90.3</b>
<b>Total Service Records Analyzed</b>	<b>4,316</b>		<b>4,997</b>		<b>4,997</b>	

**Table 2A. Percentages of Services Meeting IEP Frequency, 2003-04 (AIR) and 2004-05 (PERB and AIR), by Service**

Service	2003-04 AIR		2004-05 PERB		2004-05 AIR Analysis of PERB Data <sup>1</sup>	
	N	%	N	%	N	%
Adaptive PE	492	75.9	604	67.6	599	66.9
Deaf/Hard of Hearing	106	53.0	223	61.6	223	61.4
Language and Speech	327	48.1	401	41.6	401	41.4
LRE	14	35.0	116	80.0	116	78.9
Mental Health	62	59.6	144	45.3	144	45.0
Non-Public Agency	19	86.4	31	79.5	31	77.5
Occupational Therapy	111	43.7	202	55.0	202	56.1
Physical Therapy	30	46.2	44	61.1	44	61.1
Pre-School	n/a	n/a	41	80.4	41	82.0
RSP	22	71.0	299	60.0	292	59.1
Visual Impairment	75	48.1	188	62.5	188	61.6
<b>Total</b>	<b>1,258</b>	<b>57.2</b>	<b>2,293</b>	<b>57.2</b>	<b>2,281</b>	<b>56.8</b>
<b>Total Service Records Analyzed</b>	<b>2,200</b>		<b>4,011</b>		<b>4,016</b>	

<sup>1</sup> AIR conducted analyses on PERB data.

**Table 3A. Percentages of Services Meeting Total Duration, 2003-04 (AIR) and 2004-05 (PERB and AIR), by Service**

Service	2003-04 AIR		2004-05 PERB		2004-05 AIR Analysis of PERB Data <sup>1</sup>	
	N	%	N	%	N	%
Adaptive PE	435	69.0	593	67.3	588	66.4
Deaf/Hard of Hearing	56	62.2	230	63.7	230	63.2
Language and Speech	348	52.8	426	44.4	426	44.1
LRE	12	44.4	115	79.9	115	78.8
Mental Health	74	71.8	194	62.0	194	61.4
Non-Public Agency	22	100.0	35	83.3	35	83.3
Occupational Therapy	142	59.4	203	55.6	203	56.2
Physical Therapy	29	52.7	43	59.7	43	59.7
Pre-School	n/a	n/a	44	86.3	44	88.0
RSP	23	74.2	298	61.2	295	59.2
Visual Impairment	12	60.0	201	66.6	201	65.5
<b>Total</b>	<b>1,153</b>	<b>61.5</b>	<b>2,382</b>	<b>59.9</b>	<b>2,374</b>	<b>59.2</b>
<b>Total Service Records Analyzed</b>	<b>1,876</b>		<b>3,977</b>		<b>4,007</b>	

<sup>1</sup> AIR conducted analyses on PERB data.

**Appendix B**

**Table 1B: Numbers of First Year Site Visit Observations by Disability: Using First Observations Only and Re-Categorizing “Provider Absent” as Valid Evidence of Service Provision**

	Service Provided						Service Not Provided			Total
	Provided Total	Session Completed	Service Provided / Session Incomplete	Student Absent	Student No Show	Provider Absent	Not Provided Total	Provider in Meeting	Other	
AUT	23	16	2	0	4	1	6	5	1	29
D/HH	19	15	0	1	3	0	7	0	7	26
ED	17	15	1	0	0	1	2	1	1	19
MD/DB	30	24	1	0	5	0	2	1	1	32
MR	25	21	2	0	2	0	2	0	2	27
OHI	18	16	1	0	1	0	5	1	4	23
ORT/TBI	27	23	1	1	1	1	5	1	4	32
SLD	24	19	3	1	1	0	2	1	1	26
SLI	25	18	3	1	1	2	3	1	2	28
VI	23	17	4	0	0	2	5	2	3	28
<b>Total</b>	<b>231</b>	<b>184</b>	<b>18</b>	<b>4</b>	<b>18</b>	<b>7</b>	<b>39</b>	<b>13</b>	<b>26</b>	<b>270</b>

**Table 2B: Percentages of First Year Site Visit Observations by Disability: Using First Observations Only and Re-Categorizing “Provider Absent” as Valid Evidence of Service Provision**

	Service Provided						Service Not Provided			Total
	Provided Total	Session Completed	Service Provided / Session Incomplete	Student Absent	Student No Show	Provider Absent	Not Provided Total	Provider in Meeting	Other	
AUT	79.3%	55.2%	6.9%	0.0%	13.8%	3.4%	20.7%	17.2%	3.4%	100%
D/HH	73.1%	57.7%	0.0%	3.8%	11.5%	0.0%	26.9%	0.0%	26.9%	100%
ED	89.5%	78.9%	5.3%	0.0%	0.0%	5.3%	10.5%	5.3%	5.3%	100%
MD/DB	93.8%	75.0%	3.1%	0.0%	15.6%	0.0%	6.3%	3.1%	3.1%	100%
MR	92.6%	77.8%	7.4%	0.0%	7.4%	0.0%	7.4%	0.0%	7.4%	100%
OHI	78.3%	69.6%	4.3%	0.0%	4.3%	0.0%	21.7%	4.3%	17.4%	100%
ORT/TBI	84.4%	71.9%	3.1%	3.1%	3.1%	3.1%	15.6%	3.1%	12.5%	100%
SLD	92.3%	73.1%	11.5%	3.8%	3.8%	0.0%	7.7%	3.8%	3.8%	100%
SLI	89.3%	64.3%	10.7%	3.6%	3.6%	7.1%	10.7%	3.6%	7.1%	100%
VI	82.1%	60.7%	14.3%	0.0%	0.0%	7.1%	17.9%	7.1%	10.7%	100%
<b>Total</b>	<b>85.6%</b>	<b>68.1%</b>	<b>6.7%</b>	<b>1.5%</b>	<b>6.7%</b>	<b>2.6%</b>	<b>14.4%</b>	<b>4.8%</b>	<b>9.6%</b>	<b>100%</b>

**Appendix C**

**Table 1C. Differences in frequency compliance rates if 1 deviation from required frequency treated as being in compliance\***

	Frequency Compliance Rate with 1 Month		Frequency Compliance Rate with 8-Weeks	
	No deviations allowed	1 deviation allowed	No deviations allowed	1 deviation allowed
	%	%	%	%
APE	15.9	29.5	64.4	82.2
AUD	66.7	93.3	58.8	58.8
VI	9.1	54.5	50.0	70.0
D/HH	55.6	100.0	77.8	88.9
LRE	72.7	100.0	85.2	96.3
LAS	23.8	81.0	45.5	59.1
SMH	33.3	80.0	53.3	66.7
NPA	75.0	75.0	73.3	93.3
OT	47.1	91.2	65.7	80.0
PKIT	85.7	100.0	85.7	92.9
PT	58.3	83.3	76.9	76.9
RSP	50.0	50.0	59.1	59.1
<b>Total</b>	<b>46.0</b>	<b>73.0</b>	<b>66.0</b>	<b>77.5</b>
<b>Total Service Records Analyzed</b>	<b>237**</b>	<b>237**</b>	<b>244</b>	<b>244</b>

\* One deviation means one missed service session (in addition to allowed absences such as illness, family emergency, and jury duty) will constitute compliance with the required IEP frequency.

\*\* Total number of service records lower for 1 month analysis in comparison to the 8-week analysis, as some records may be missing October or January logs