



WAYNE L. WELSH, CPA
AUDITOR GENERAL

STATE OF UTAH

Office of the Legislative Auditor General

130 STATE CAPITOL • PO BOX 140151 • SALT LAKE CITY UT 84114-0151
(801) 538-1033 • FAX (801) 538-1063

Audit Subcommittee of the Legislative Management Committee

President R. Lane Beattie, Co-Chair • Speaker Martin R. Stephens, Co-Chair
Senator Mike Dmitrich • Representative Brad King

December 16, 1999

Members of the Human Services Interim Committee
Members of the Child Welfare Oversight Committee
Utah State Capitol Building
Salt Lake City, UT 84114

Subject: Bureau of Services Review (Report #99-11)

Dear Legislators:

In accordance with **Utah Code 62A-4a-118**, our office has conducted its annual audit of the Bureau of Services Review's (BSR) examination of the Division of Child and Family Service's (DCFS) child welfare casework. This report builds on our audit of BSR's 1998 review (Report #99-05). BSR did not have the opportunity to implement the recommendations from that audit before case reviews were conducted for its 1999 review of DCFS. Because of this, we limited the scope of our audit work reported here. Primarily, we update our disagreement rate with BSR for 1999 cases and suggest some strategies to reduce the number of cases BSR reviews.

BSR was created to annually review a sample of DCFS cases and report "...findings regarding whether state statutes, division policy, and legislative policy were followed in each sample case." Since its inception, BSR has focused primarily on measuring compliance to detailed process requirements. Beginning with our audit in 1995, our reports have recommended that BSR create a more comprehensive review that also measures caseworker performance and outcomes. BSR is in the process of implementing a more comprehensive review.

BSR is currently implementing a qualitative review as part of the court-ordered Performance Milestone Plan for Utah's child welfare system. The court also appointed a consulting group to be the court monitors of the new plan. Toward this end, the consulting group and BSR created a monitoring and review system that includes a reduced case process review and a new qualitative review that evaluates outcomes and quality of casework practice. The new monitoring system is currently being implemented, and the results will be revealed in BSR's 2000 report. Accordingly, our office will review those results next year.

The three main sections of this report cover the following topics:

- Recommendations from our audit of BSR's 1998 review remain valid and should be implemented in its 2000 review.
- BSR's 1999 review of DCFS compliance with case process requirements is generally accurate. In fact, BSR's accuracy is somewhat better than last year.
- BSR samples more case files than needed, leading to additional reviewer errors. BSR could follow a number of strategies to reduce its sample size.

Prior Recommendations Have Yet to Be Implemented

Our current audit is based on BSR's 1999 review that was released in September 1999. Similar to the prior year, the 1999 review focused on measuring compliance to the now defunct settlement agreement. As mentioned earlier, BSR did not have the opportunity to implement the recommendations from our audit of its 1998 review before the 1999 case reviews were conducted. As a result, the recommendations from our prior report remain valid and are briefly reviewed in this section. BSR officials have informed us that they are implementing these recommendations.

1. ***BSR Should Include an Overall Assessment of Decisions and Casework Quality on Each Case it Reviews.*** A significant weakness of compliance monitoring has been its failure to reach overall conclusions regarding decisions and casework quality. The case process review instrument asks questions concerning the documentation of individual requirements in a case, but not about the overall appropriateness of

actions taken. Our audits have found that the compliance scores do not necessarily relate to how well DCFS protected a child or preserved a family. BSR should reach an overall conclusion of casework quality each time a case file is reviewed.

2. ***BSR Should Monitor and Evaluate the Intake Process.*** BSR chose not to review the intake process and unaccepted referrals for its 1999 review. This decision is a major concern because of the significant problems we discovered and reported in our last audit. BSR should monitor the intake process and unaccepted referrals because of the risks associated with improper decisions by DCFS intake workers. When a referral is not accepted, the child involved receives no protection from the state. BSR officials have told us that they will review the intake process beginning next year.
3. ***A More Comprehensive Review Process Is Needed.*** A broader, more comprehensive case review process which evaluates overall caseworker performance and outcomes, and utilizes multiple sources of information beyond the case files, can more accurately reflect system effectiveness. For example, personal interviews could be conducted with appropriate individuals involved with a case. As noted earlier, our audits have recommended this course of action since 1995. BSR and the consulting group are currently implementing a new monitoring and review system that includes a qualitative review.
4. ***BSR Should Focus on Accuracy Rather than the Number of Cases Reviewed.*** We feel that BSR samples an unnecessarily large number of case files for its reviews. In our opinion, this large sample size results in a high number of reviewer errors. BSR has reduced its sample size, but we feel that it could be reduced even further without compromising statistical reliability. In the last section of this report we describe some strategies BSR should consider to reduce its sample size.
5. ***BSR Should Refine and Reduce the Number of Questions it Reviews.*** The questions used to review for compliance should be limited to the most significant aspects of casework that lend themselves to compliance reviews. Questions that attempt to measure broad, qualitative concepts are better dealt with as either part of a quality review or specifically as performance-based questions. Reducing the number of questions will increase the efficiency of the reviews. BSR made reductions in its 1999 review, but we feel further reductions are possible.

6. *BSR Should Continually Refine the Case Reviewer Manual.* We feel that this process is necessary and valuable because it provides updated training and clarity for reviewers. It may also prevent overuse of the “Not Applicable” scoring option. During the case reviews of our current audit, we noted one set of questions that was being scored differently by a number of the BSR reviewers. When we made inquiries about the differences, we discovered that the reviewers had different interpretations of what the questions were asking. This situation resulted in several inaccurate reviews by BSR.
7. *BSR Should Continue to Formalize a Double-reading Process.* This process consists of a small number of case files being analyzed and scored a second time by a different BSR reviewer. The answers of each are then compared to help ensure consistency among reviewers and also to identify problematic or confusing questions. BSR has made changes to its double-reading process since our last audit, and we will analyze those changes next year.

BSR Is Generally Accurate in Rating DCFS Compliance

As we found in our previous audits, our current findings show that BSR is generally accurate in rating DCFS compliance with procedural requirements. Our scoring methodology remained consistent with our past audits, although there were some changes in our sampling methodology. We found that BSR improved its accuracy from our last audit in rating DCFS compliance. In addition, we address the different types of disagreements we had with BSR reviewers and make suggestions that may improve future case process reviews.

Our Scoring Methodology Remained Consistent with Prior Audits. In following the methodology of our past audits, our current audit used the BSR requirements to analyze and score a sub-sample of the cases previously reviewed by BSR. This process results in an error rate which represents our disagreements with BSR scoring. All sampled case files were reviewed and scored without knowing how the original BSR reviewer had scored each requirement. All of our disagreements were first discussed as an audit team. We then discussed each disagreement with one of the permanent BSR reviewers or members of BSR management. In some instances, additional BSR reviewers or members

of management were consulted. The purpose of consulting with BSR was to thoroughly understand the requirements and the causes of the disagreements.

Some changes were made by our audit team in an effort to increase the efficiency of our audit. In doing so, we limited our sample to CPS General Investigation cases, PSS court-ordered home-based services cases, and PSC voluntary home-based services cases. This provides an analysis of services for a family that is entering the child welfare system, as well as an analysis of services once a family is within the system. We also limited our audit to cases selected from the three Salt Lake regions, and from the Richfield and Cedar City offices of the Southwest Region. This saved time in locating case files and gave our audit a roughly proportionate sample of cases from urban and non-urban DCFS offices throughout the state.

Disagreement Rate Is Generally Consistent with Past Audits

We disagreed with the scoring on 9 percent of the compliance items that we analyzed from BSR's 1999 case process review. This is generally consistent with the disagreement rates of our prior audits of 1995 (9%), 1996 (7%), and 1998 (12%). The results of our current analysis, in terms of disagreements with BSR scoring, are shown in Figure 1.

Figure 1. Disagreement Rate for BSR's 1999 Review of DCFS. BSR is generally accurate in rating DCFS compliance.

Type of Case Review	Number of Questions Reviewed	Number of Disagreements with BSR Scoring	Disagreement Percentage
CPS General Investigation	382	29	7.6%
Home-based PSS and PSC	<u>118</u>	<u>15</u>	<u>12.7</u>
Totals	500	44	8.8%

We classified only two of the 44 disagreements (0.4%) as potentially significant in terms of child safety or well-being. This is an improvement from 1.6 percent in 1998. We also

found no evidence or pattern suggesting that BSR is too lenient in reviewing DCFS casework. The overall ratings of DCFS casework compliance were similar with BSR's at 83 percent and our's at 81 percent.

Disagreements Are Caused by Various Factors. Our audit team categorized each of the disagreements into one of four types: reviewer error, differences in judgement, untimely DCFS documentation, and objections with BSR policy. While some disagreements are unavoidable in this type of subjective review process, there are some steps BSR can take to reduce its errors.

- **BSR Reviewer Errors Accounted for 50% of Our Disagreements.** A reviewer error occurs when either the BSR reviewer missed documentation in the case file that indicated DCFS compliance or non-compliance with the requirement, or the reviewer did not answer according to the instructions in the case reviewer manual. In discussing our disagreements with BSR staff, we were occasionally told that documentation in the files could have been missed by the original BSR reviewer because cases are reviewed quickly. We believe that some of these errors could be eliminated if the reviewers had time to more thoroughly review each case. BSR could improve in this area by focusing on accuracy rather than the number of cases reviewed. The next section of this audit addresses strategies to reduce the number of cases sampled. BSR could also improve in this area by reducing the number of questions it reviews. In addition, BSR should continue to use its double-reading process to identify reviewer errors that might be prevented by better training or improved questions.
- **Difference in Judgement Accounted for 27% of Our Disagreements.** A judgment disagreement occurs when there is a legitimate difference in interpretation of information between reviewers. More than one answer may have some merit depending on the circumstances. Some judgement disagreements are inevitable in this type of subjective review process. However, an effective double-reading process helps BSR minimize judgement disagreements by identifying questions that are too subjective and areas where the case reviewer manual may be improved.
- **Untimely DCFS Documentation Accounted for 18% of Our Disagreements.** Our audits are conducted several months after BSR concludes its reviews. Documentation can be added or removed from the case files during that time period. We were careful to note dates on documents when reviewing the case files. In doing

so, we found that the caseworkers occasionally completed their activity logs after BSR had completed its review. The activity logs are vital for accurate case file reviews. While these situations are certainly not the fault of BSR reviewers, we did note them as system errors because the lack of information resulted in inaccurate reviews. DCFS supervisors should ensure that information is completed and placed in the case files in a timely manner.

- **Our Objections with BSR Policy Accounted for 5% of Our Disagreements.** Objections occur when the BSR reviewer answered correctly according to the case reviewer manual, but our audit team did not agree with the instructions and/or the BSR policy. All of these objections resulted from what we felt were inappropriate uses of the “Not Applicable” scoring option. BSR could improve in this area by rewording questions or refining the case reviewer manual to minimize use of the “Not Applicable” scoring option.

BSR Could Employ a Variety of Strategies to Reduce Sample Size

We believe that BSR should reduce the number of cases it samples in order to reduce reviewer errors and free-up resources for other activities. In our last report, we concluded that BSR should focus on accuracy rather than the number of cases reviewed. While BSR did reduce its sample size by 15 percent from the previous year, we feel that further reductions are possible. This section discusses some strategies BSR should consider to further reduce the sample size for its 2000 review.

State statute requires BSR to take a statistically significant sample of cases, but we believe BSR can do that with a much smaller sample size. Two types of factors determine the sample size required for statistical significance. First are the characteristics of the population itself, including its error rate, size, and how it is divided into sub-groups. Second are parameters chosen by evaluators based on their judgement: confidence level and precision range. We developed three strategies that BSR could use to reduce its sample size based on population characteristics. Figure 2 shows that incorporating all three strategies could reduce BSR’s sample size by more than half.

Figure 2. Our Suggested Sample Size Compared to BSR’s 1999 Sample. BSR could reduce its sample size by more than half, thereby increasing efficiency and reducing reviewer errors.

Case Type	Population	1999 BSR Sample Size	Suggested Sample Size*	Percent Reduction
CPS	4337	584	188	68%
Unable to Locate	101	88	67	24
Foster Care	2632	298	182	39
PSS & PSC	1843	480	177	63
PFP	292	149	117	21
Foster Care Re-entry	<u>207</u>	<u>127</u>	<u>101</u>	<u>20</u>
Totals	9412	1726	832	52%

* Based on 95% confidence level, 5% precision, and 15% error rate.

The purpose of a review helps determine the most appropriate way to control the sample size. Ultimately, the most appropriate sampling plan depends on what information the review is designed to produce. We did not have time to thoroughly evaluate the implications of the strategies discussed here, but we think BSR should consider these with other strategies it develops to limit its sample size. The three strategies we suggest are:

- Base Expected Error Rate on Performance Goals.** BSR could reduce its sample sizes by basing them on performance goals rather than historical performance levels. Precision would not be significantly affected by reducing the sample sizes. We calculated sample sizes assuming error rates of 15 percent, because that is the highest acceptable error rate in the milestone plan. For example, we calculated a foster care sample size of 182. In comparison, BSR’s larger sample size of 298 is needed assuming an error rate of 32 percent. We believe that if actual performance is that poor, then it probably is not necessary to measure with 5 percent precision. In fact, if an error rate of 32 percent is found with a sample size of 182, the statistical precision is only increased to 6.5 percent.

- **Combine Sub-groups When Possible.** Dividing populations into sub-groups greatly increases sample size. In reviewing two sub-groups that BSR sampled independently, we found that the review questions were identical and the results were quite similar. Therefore, in our audit work, we took one sample from PSS and PSC home-based services cases combined rather than taking two separate samples. If BSR sampled from the two sub-groups combined, it would practically cut in half the number of these types of cases reviewed.
- **Reconsider How CPS Sub-groups Are Included.** We believe that BSR should reconsider its sampling and reporting of CPS cases. BSR takes a random sample of all CPS cases and then adds most or all of three sub-group types: priority 1, medical neglect, and shelter care. These sub-groups are added because some review questions apply specifically to them, but the case types are so rare that they infrequently appear in the random sample. Thus, of the 584 CPS cases sampled, about half are from the sub-groups even though they make up only about 7 percent of the population. Nevertheless, BSR has reported its results without regard for the over-representation of the sub-groups. Depending on a more in-depth analysis of its data, we think there are two ways BSR could consider handling the sub-groups differently:
 - a. If the cases in the sub-groups are representative of CPS General Investigation cases, then the results of all cases from a smaller sample can be combined. BSR staff helped us complete a cursory review of the 1999 CPS results with and without the sub-groups included. The results appeared to be similar, indicating that the sub-group cases are representative of CPS General Investigation cases. If this is so, then one way to reduce sample size is to first sample the sub-groups and then supplement them with additional CPS General Investigation cases as needed. This strategy, which is reflected in the sample size we show in Figure 2, is consistent with BSR's practice of combining the results from the sub-groups and CPS General Investigation cases.
 - b. If in-depth analysis shows that results of the sub-groups are dissimilar from the CPS General Investigation cases, then the sub-groups should not be combined. In this situation, we believe that BSR should sample the sub-groups separately, but only review and score those specific review questions that are applicable. BSR should then report the results of the sub-groups separately from the CPS General Investigation questions.

Members of the Human Services Interim Committee
Members of the Child Welfare Oversight Committee
December 16, 1999
Page 10

BSR needs to consider these strategies as well as other ways of reducing the sample size. BSR officials stated that for the 2000 review they are considering reducing the confidence level to 90 percent to reduce the sample size. BSR is also going to reduce the number of questions to decrease the amount of time spent reviewing the cases. BSR should use the time saved by reducing the sample size to focus on the accuracy of the case process review and other important activities such as implementing the qualitative review.

Recommendations:

2. We recommend that BSR implement the recommendations from our audit of BSR's 1998 review for its 2000 review.
3. We recommend that BSR reduce its sample size by implementing the strategies described in this report.

We hope this letter gives you the information you need. A response from the Department of Human Services is attached. If you would like additional information or further clarification of any point, please feel free to contact our office.

Sincerely,

Wayne L. Welsh, CPA
Auditor General

WLW:GRF/lm



State of Utah

DEPARTMENT OF HUMAN SERVICES
OFFICE OF THE EXECUTIVE DIRECTOR
www.dhs.state.ut.us

120 North 200 West #319
Salt Lake City, Utah 84103
(801) 538-4001
(801) 538-4016 (FAX)
dirdhs@email.state.ut.us (E-mail)

Michael O. Leavitt
Governor
Robin Arnold-Williams
Executive Director
Douglas E. West
Deputy Director

December 8, 1999

Wayne L. Welsh, Auditor General
Office of the Legislative Auditor General
130 State Capitol
Salt Lake City, UT 84114-0151

Dear Mr. Welsh:

We believe the Legislative Auditors Office has done a thorough evaluation of the accuracy of case file reads. As noted in our response to their audit of 1998 data, we were anxious for them to evaluate whether some steps we have taken resulted in more accurate reads. The auditors evaluation shows that accuracy has improved significantly (auditors disagreed with 12 percent of our responses in the 1998 review; auditors disagreed with 8.8 percent of our responses in the 1999 review). We will continue to work to reduce this number further. We note that the disagreement rate would be 7.2 percent if the untimely documentation were reported separately. In these cases, the logs were not present when BSR reviewed the file but were put in later. Consequently, the Auditor does not disagree with the BSR readers scoring, but rather is pointing out a system weakness.

Concerning last year's recommendations, as we explained to the auditors, we are in process of implementing many of the recommendations they made in their July 1999 report. For example, we are including intake questions as part of our case process review as well as including several qualitative-type questions on our CPS review.

As far as reducing sample sizes, we will evaluate the strategies suggested in the report. As noted, we have already substantially reduced the number of questions reviewed as well as the number of cases sampled. Auditors make several additional suggestions on how we can sample fewer cases while still maintaining acceptable reliability. These are good suggestions. We will carefully review them and select those we feel will be most beneficial.

Thank you for your assistance in these areas.

Sincerely,

A handwritten signature in cursive script that reads "Robin Arnold-Williams".

Robin Arnold-Williams, D.S.W.
Executive Director