

## **Data Validation Policy**

### **Reference:**

The Workforce Investment Act (WIA) of 1998; Federal Register/Vol. 65, No. 156/Friday, August 11, 2000; TEGL 3-03; TEGL 9-07; TEN 14-02; TEGL 27-10, Change 1; TEGL 28-11; WIA Data Reporting and Validation System User Handbook; and Employment and Training Administration (ETA) Presentation.

### **Background:**

In 2001, the President announced a Management Agenda to improve the management and performance of the federal government. It emphasized the importance of complete information for program monitoring and improving program results. This expectation concurs with wording in Sec. 185(c) of the Act: "Each State, each local board, and each recipient (other than a subrecipient, subgrantee, or contractor of a recipient) receiving funds under this title...shall prescribe and maintain comparable management information systems, in accordance with guidelines that shall be prescribed by the Secretary, designed to facilitate the uniform compilation, cross tabulation, and analysis of programmatic, participant, and financial data, on statewide, local area, and other appropriate bases, necessary for reporting, monitoring, and evaluating purposes, including data necessary to comply with section 188."

Performance audits conducted by the U.S. Department of Labor's Office of Inspector General found that the accuracy of reported performance outcomes could not be assured due to insufficient local, state, and federal oversight. To address these concerns, and to ensure the accuracy of data collected and reported on the workforce investment system, the Employment and Training Administration (ETA) expanded its efforts to conduct data validation. During PY 2003 Nebraska implemented data validation for the WIA Title IB programs. Attached to this policy is a revised list reflecting the source documentation requirements for PY 2012 WIA data element validation.

### **Action:**

After the 10 day public review period, this policy is considered final. All programs utilizing NEworks must adhere to this policy. Program management should inform all NEworks end users of this policy. Questions and comments should be submitted in writing to Jan Fox, Policy Coordinator, [jan.fox@nebraska.gov](mailto:jan.fox@nebraska.gov).

### **Policy:**

The data validation initiative covers both the accuracy of aggregate reports submitted to ETA on program activity and performance outcomes and the accuracy of individual data elements. For PY 2012, Nebraska must submit the Quarter 4 WIA Standardized Record Data (WIASRD) files by September 15, 2013, and a validated WIA Annual Report by close of business (COB), October 1, 2013. Report validation must be completed prior to submitting the annual report tables, and data element validation using a file of exiters and participants reported in the annual report must be submitted by February 1, 2014.

### **Report Validation**

This process evaluates the validity of aggregate reports submitted to ETA by checking the accuracy of the reporting software used to calculate the reports. Report validation is accomplished by processing an entire file of participant records into validation counts and comparing the validation counts to those reported by the State or grantee.

ETA has developed standardized software that can be used to validate the accuracy of aggregate reports generated by State or grantee software. The latest standardized software and user handbooks for each program are posted on ETA's web site at:

[http://www.doleta.gov/performance/reporting/tools\\_datavalidation.cfm](http://www.doleta.gov/performance/reporting/tools_datavalidation.cfm). The ETA software will produce an error rate for each reported count. The standardized software can also be used to generate aggregate information to be submitted to ETA prior to submitting the annual report.

### **Data Element Validation Process**

This process assesses the accuracy of participant data records. Data element validation is performed by reviewing samples of participant records against source documentation to ensure compliance with federal definitions.

Not all data elements are subject to validation. Data elements are selected for validation based on three factors:

- **Feasibility** – ETA can validate data elements only when it is practical and efficient to locate and examine supporting evidence within the state records. Therefore, such items as race, ethnicity, and gender will not be validated because these data elements are self-reported by participants, and it is not efficient to locate the participant to document these characteristics. It is also not practical to validate for data entry errors.
- **Risk** – The process for validating data elements is based partly on the likelihood that the data element may be inaccurate. Data elements involving human judgment are more prone to error than data elements that do not involve human judgment. For example, determination of employment based on supplemental sources is more likely to be in error than determination of employment from wage records.
- **Importance** – Data elements are selected for validation based primarily on their importance to the integrity of the ETA 9091 report (Annual Report).

The ETA standardized software generates a sample of the participant records and data elements the State or grantee will validate. Data element validation samples are selected from a sample of offices within the State, rather than from every office.

Two sampling techniques are used:

1. **Clustering** – The software selects the sample of exiters from a relatively small number of locations within the State where source documentation is stored. This approach reduces the burden that on-site validation imposes on the State by limiting the number of locations to visit.
2. **Stratification** – Cases that have a higher potential for error that will affect performance (positive employment or education outcomes) and more data elements to be validated are sampled in greater proportion than cases with a reduced chance of error (negative employment or education outcomes) and fewer elements to validate.

Once the State has validated the sampled records, the software weights the results of the validation to correct for over- and under-sampling resulting from the clustering and stratification.

The data element validation tasks include:

1. **Assemble worksheets and arrange visits to Career Center offices** – If possible, State staff should conduct the validation on site except for data stored at the state level, i.e., wage record information. States have the flexibility to determine the most efficient source to use to validate the information when more than one source is specified. State staff shall notify local staff well in advance of when the onsite validation will occur to ensure local staff is available to assist in the validation. State staff shall also inform local staff which records have been sampled, no more than one to two days in advance of the onsite review. This will make the review more efficient, while minimizing the possibility for changes to the case files.
2. **Assemble supporting documentation and complete worksheets** – The attached list identifies acceptable sources of information to meet documentation requirements.
3. **Examine summary and analytical reports** – The software calculates the error rate for each element and produces separate summary and analytical reports for each group.

To ensure integrity and authority of the data validation process, Nebraska conducts the data validation process through the Nebraska Department of Labor's Office of Legal Counsel. The staff monitor who is assigned to this office is responsible for conducting the review of participant records, and the submittal of the annual data validation report to ETA. Participant records will be reviewed in both paper and electronic format.

### **Accuracy Standards**

According to TEGL 9-07, ETA considers that a WIA State grantee fails to report year-end results when the grantee fails to submit an accurate WIA annual report as determined by applying a two percent error threshold to program outcomes, including numerators and denominators, summarized in the Report Validation Summary.

States and grantees will be held accountable for meeting acceptable levels and will be required to address any issues concerning data accuracy. States and grantees that fail to meet accuracy standards will receive technical assistance from ETA and will develop and implement a corrective action plan. Data that do not meet accuracy standards will not be acceptable for measuring performance, and may keep the State or grantee from being eligible for incentives that are awarded based on performance data.

Upon completion of data validation in Nebraska, the results will be shared with the local WIA area. In incidents where acceptable levels are not met, technical assistance may be provided. Depending on the degree of deficiency, the State may require a corrective action plan from the local area.

### **Data Elements and Source Documentation**

The elements required for data element validation focus on eligibility, performance, and services. Attached is a list providing federal guidance regarding what types of source documentation are acceptable to validate key data elements related to WIA program eligibility, services and outcomes. Only one of the documents listed for each element is required for validation. However, if multiple sources are collected for the same data element and the sources conflict, the most reliable source should be used to determine if the element passes or fails. For example, for School Status at Participation (DRVS Field #30) copies of records from an educational institution are a more reliable source than participant self-attestation. Special attention must be paid to the following general definitions for certain types of source documentation:

- **Cross-Match:** A cross-match requires validators to find detailed supporting evidence for the data element. An indicator or presence of an SSN in a non-WIA database is not

sufficient evidence. For example, Temporary Assistance to Needy Families (TANF) participation can be determined by a cross-match with Nebraska's public assistance database. It is not sufficient to find that the sampled SSN is present in the public assistance database; validators must also find supporting information, such as, dates of participation and services rendered.

- **State MIS:** Unless otherwise noted, State MIS refers to specific, detailed information that is stored in the State's information system that supports an element. An indicator, such as a checkmark on a computer screen, is not acceptable source documentation in and of itself. For example, State MIS is an acceptable source documentation for date of first training service. To be an acceptable source to validate the date of first training service, in addition to the date of first training, the State MIS should have information about the type of training and the organization that provided the training. This detailed information makes valid source documentation and makes it unnecessary for such states to validate this data element in local offices.
- **Self-Attestation:** Self-attestation occurs when a participant states his or her status for a particular data element, such as pregnant or parenting youth, and then signs and dates a form acknowledging this status. The key elements for self-attestation are:
  - the participant identifying his or her status for permitted elements and
  - signing and dating a form attesting to this self-identification.

The form and signature can be on paper or in the State MIS, with an online signature.

- **Case Notes:** Case notes refer to either paper or electronic statements by the case manager that identifies, at a minimum, the following:
  - a participant's status for a specific data element,
  - the date on which the information was obtained, and
  - the case manager who obtained the information.

The WIA Data Reporting and Validation System User Handbook provides the validation rules that apply to each data element, i.e., match or support. To match, the data on the worksheet must be the same as the data in the source documentation. For example, if the worksheet says a participant's date of birth is July 1, 1975, then the source documentation must also have July 1, 1975 as the birth date. To support the data on the worksheet, the source documentation must provide evidence that the data on the worksheet is correct. This instruction is used when information must be interpreted or processed before it can be used to assess the accuracy of the data on the participant's records. For example, source documentation is used to support youth who need additional assistance because validators must interpret policy and determine if the documentation supports that policy.

### **Record Retention**

The record retention requirement that will apply to records of all validation activities, including files, worksheets, reports, and source documentation, is three years after completion. The State shall be responsible for keeping completed worksheets, summary reports and analytical reports. Local offices/Workforce Investment Boards shall retain the validated files and original supporting documentation for monitoring purposes.

### **Source Documentation Requirements for PY 2012 WIA Data Element Validation**

The U.S. DOL will not issue the annual Reporting Timelines Training and Employment Guidance Letter (TEGL) for PY 2012 due to some of the uncertainty created by the launch of the Enterprise Data Reporting and Validation System (E-DRVS). TEGL 28-11 will serve as an appropriate guide for reporting timelines, as many of the dates therein are the same for PY 2012.