Departmental Response to Comments Received from the 60-Day Federal Register Notice on Data Validation

	STATE COMMENTS/ISSUES	ETA RESPONSE	STATES PROVIDING COMMENTS
	REVIEW FOCUS 1: Evaluate whether the proposed collection of information is necessary for the proper performance of the functions of the agency, including whether the information will have practical utility.		
1	Data validation activities have value and the states generally support the Department's efforts to promote and refine the requirements. One state commented that the state used data validation to improve the quality of program and performance data over the past several program years and another commented that it learned there is a need for more policy guidance from the state to the local level regarding case management processes and file management.	The Department is encouraged by these observations and comments, as none of the commenters took issue with the overall concept and value of data validation. Since the implementation of data validation, states are paying greater attention to how data is collected and what information is being used to establish priority of service and calculate performance outcomes. States should use the data validation guidelines as a basis for standardizing data collection and record keeping across local areas to ensure data quality across all workforce development programs.	WI, IL, OR
	REVIEW FOCUS 2: Evaluate the accuracy of the agency's estimate of the burden of the proposed collection of information, including the validity of the methodology and assumptions used.		
2	The burden associated with data validation is underestimated.	As indicated in section 12.A of the Supporting Statement, the overall estimate for state programs is based on averaging separate estimates for small, medium, and large states. As shown in Table 2 of section 12.A, the estimated burden for large states such as New York, Oregon and Pennsylvania is 1,206 hours,	PA, WI, IL, OR, NY

while the estimated burden for medium states, such as Illinois and Wisconsin is 746 hours. The corresponding cost estimate for large and medium states (not shown in the table) is $39,195 and $24,245 respectively, calculated by dividing the total cost by the number of states. This methodology was used to acknowledge that there would be more burden on larger states, due to the sample sizes required to make the data validation statistically valid, which is balanced by the fact that larger states generally have more resources to conduct data validation than smaller states. However, there will be expected variations in burden within each subcategory (small, medium, and large states), as factors such as the sophistication of case management processes and systems and geographic size of the state affect individual state burdens.

Wisconsin further stated that it believes that "start-up" costs are still being incurred for validation that were not included in the burden estimate in view of the time and effort required to upload and install software updates and associated patches. Illinois also expressed concern that an additional burden is incurred because of the time and effort required to install the updated software releases, to make data adjustments for changes in the software requirements, as well as the staff training and time pressures associated with meeting validation deadlines using the new releases, none of which were addressed in the estimates. The Department concedes that these appear to be valid concerns, as there have been on-going difficulties in

#	Comment	Response	
		the Data Reporting and Validation Software (DRVS) development process. The Department is actively trying to address these problems to lessen the associated burden. In the meantime if a software release is delayed, the Department has generally given states more time to complete the requirement to alleviate time pressures and has tried to ensure that the record layout and edits have generally remained constant throughout the software upgrade process.	
3	There is a need to develop common data element validation policies/requirements across the USDOL-administered programs listed in the proposed Information Collection Request (ICR), rather than "silo" requirements for each program.	While none of the respondents provided enough/complete information to allow the Department to revise or question the general accuracy of its original burden estimates, the Department will explore ways to lessen the burden on states associated with data validation by exploring policy changes that reduce the amount of travel required for data element validation, or the frequency of data validation, with the Office of the Inspector General. The Department generally agrees with this comment and has made efforts to do so to the extent possible, within the constraints of the currently approved OMB performance reports for these programs and the different legislations that authorize the programs. The Department plans to take additional steps in this direction for the state-based programs with the development of the data validation requirements associated with the implementation of the Workforce Investment Streamlined Performance Reporting System (WISPR), the proposed consolidated reporting system for the Wagner-Peyser/Veterans Employment	NY

		and Training Service (VETS) funded programs, the Workforce Investment Act (WIA) Title I-B programs, and the Trade Adjustment Assistance (TAA) program. Having all of these programs using the same reporting forms and record layouts will significantly help the Department align the requirements for data element validation across these programs to ensure the elements validated and the required supporting documentation are consistent across programs.	
4	It isn't cost effective for validation teams to travel to local offices to spend time reviewing and validating participant case files.	The Department is aware that in the current atmosphere of budget cuts and rising travel expenses (particularly fuel); it is becoming less cost-effective to travel to local offices to review participant case files. However, it is necessary to review participant case files to provide some assurances that the source data that is being used to calculate the states performance reports is reasonably valid. As such, states are encouraged to combine these reviews with monitoring visits to local areas to leverage travel. Further, the Department has tried to address these concerns to some extent in the design of the software and the associated user documentation. Clustered sampling was built into the software to minimize the number of local offices visited and, in cases where it is not feasible to perform an onsite review in all locations; state staff may ship selected files to a more convenient location for review, upon approval by the Department (page 75 of the DRVS 6.3 WIA handbook). As stated earlier, the Department will explore ways to lessen the burden on states associated with data	WI, OR

	Comment	Response	State
		validation by exploring policy changes that reduce the amount of travel required for data element validation, or the frequency of data validation, with the Office of the Inspector General.	
5	The requirement to notify local staff of the records which have been sampled in advance of the on-site review should be extended from 1-2 days prior to 5 days prior.	The Department recommends limiting the amount of time local staff has to pull sampled records in order to prevent the tampering of documentation in the file. However, if a state is able to articulate a compelling reason as to why 5 days is necessary for a local area to gather files, a formal request can be submitted to the appropriate Regional office for consideration.	WI, OR
6	We hope that USDOL will consider the utility of data-sharing agreements that states have entered into with other Federal and state data collections systems to ensure the validity and accuracy of our participant and program data at the point of entry into our workforce system.	The Department assumes this comment refers to the State's ability to cross check certain data elements through other state databases. Wisconsin has negotiated the ability to cross match against the Social Security Administration's (SSA's) database to verify 4 specific data elements: Participant Name, Gender, Data of Birth and Social Security Number (SSN). If the cross-match does not match on all 4 elements, the State does not accept the SSN as "validated" and performs a review of the local file to determine if the SSN is valid. If the case manager has failed to document the SSN with acceptable sources as required in the DRVS manual, then the State marks the element as "Failed" for validation purposes. The Department supports such collaboration between agencies and encourages more states to implement this type of cross check system to assist with data validation across workforce development programs, as long as it is properly documented and implemented.	WI
7	The commenting states question the statistical validity of the	Unlike the data element validation process for the WIA	IL, OR

sample taken for data element validation for the Wagner-Peyser Employment Service (ES) program.

programs, the primary purpose of the Wagner-Peyser Employment Service data element validation sample is to determine if the extract file was constructed properly, not to calculate estimates of the error rate. There are two reasons for this:

1. Because the Wagner-Peyser Employment Service has a very large number of participants (the largest state has over 3.5 million ES records), providing statistically meaningful results would require a large sample size. The burden to validate this sample would be quite high and is not justified given that it is a universal access program (there are no eligibility requirements to participate in the program). In fact, one of the commenting states (Illinois) acknowledged that it would be hard pressed to review extra cases if the sample size was increased.

2. The Wagner-Peyser Employment Service does not require states to collect source documentation both because the documentation burden would be quite high as a result of the number of participants and because the nature of the services – many services can be accessed via the Internet – makes it difficult to collect documentation from participants. Thus, the only "source" is the state's management information system (MIS).

Consequently, the purpose of the data element validation for the Wagner-Peyser Employment Service is to determine if the routine to develop the extract files

| 8 | The commenting state contends that there is a disparity regarding the samples pulled from Local Workforce Investment Areas (LWIA). LWIAs with a larger number of exiters have far fewer records being sampled compared to smaller LWIAs. | correctly drew the data (i.e., Are the wages correct for the quarter and person sampled? Does the state's extract routine put the correct date in the correct location in the file?) Since the extract files are developed through programmed routines, the same error tends to appear in all of the files/records. The current sample size should be sufficient to uncover these types of problems.

Many factors impact how records are sampled: the relative importance of the record to the performance outcomes, the number of exiters in the sample frame, and the geographic distribution of exiters across the state.

The Department believes that its sampling methods are optimized to be statistically valid, while giving states the flexibility to decide how they want to utilize the sampling methodology programmed into the software. This decision will impact how the sample is drawn and distributed.

In the case of Oregon, the Department believes, that Oregon chose to cluster its sample by office, based on the example provided in their comments. If the state choose to draw its sample *by office*, then comparing the number of records sampled by LWIA is not the appropriate comparison, the number of records sampled by office is the appropriate comparison. For example, if a large LWIA has 10,000 exiters and 50 offices and a small LWIA has 2,000 exiters and 2 offices, the small LWIA's offices would likely have | OR |

9	Data element validation should be paperless; having hard copies of documentation is cumbersome and expensive. New York suggests that their efforts to establish a paperless staff verification system represents an innovative approach to eliminating unnecessary bureaucratic paper and should be embraced as forward thinking based on its rationale/merits.	more than double the number of exiters of the average office in the large LWIA. As a result, records in the small LWIA's offices are more likely to be selected by the "office" sampling methodology, in which case a smaller LWIA could end up having more records sampled than a larger one. If Oregon had chosen to sample *by record*, then they could expect the distribution of records by LWIA to correspond to the relative sizes of their LWIAs. While the Department supports efforts to reduce the amount of hard copy documents/paperwork required to effectively administer its programs, a system that has no tangible source documentation defeats the intent of data validation and makes a states results, both positive and negative, unverifiable. States currently have the option to scan or make digital copies of the required documentation and append them to electronic case files to eliminate the need for paper copies/make the system "paperless"; however, very few states are currently taking advantage of this option.	NY, WI
	REVIEW FOCUS 3: Enhance the quality, utility, and clarity of the information to be collected.		
10	Part of the problem (with data validation) is the Department's tendency to continuously change the performance standards and program reporting requirements faster than most state and local reporting systems can adjust to the changes.	While this comment likely had merit a few years ago, when the Department was undertaking significant efforts to reform its programs' reporting systems to capture performance outcomes and then made further revisions to implement the common performance measures (per OMB's directions), these concerns are	WI

		now less valid. The last major changes to performance calculations and data elements occurred at the beginning of Program Year (PY) 2005 (July 1, 2005), to implement the new common performance measures policy. As such, there have not been significant revisions to the requirements/policy for report validation or data element validation for more than two years.	
11	Data validation handbooks need to be updated and offer better information about what is acceptable documentation. Information should be thoroughly reviewed and streamlined to assure clarity and consistency in interpretation.	The data validation handbook is routinely updated to reflect changes to the software and provide clarification to existing definitions and requirements as requested by the states. The Department created the handbooks as a tool for states to use during validation and has encouraged all states to create state specific policies, based on Federal validation requirements, to provide a more detailed and specific context for each state's workforce system. The Department issued an updated handbook for the Wagner-Peyser (ES) and WIA programs with the recent release of Version 7.0 of the DRVS. In addition, the Department will take the opportunity to revise the handbook to ensure the elements validated and the required supporting documentation are consistent across programs when data validation is implemented for the WISPR system, as discussed previously.	NJ, NY, PA, IL, WI
12	The validation rule should be that states must have documentation that supports the validity of a specific element, not matches it. One example is the validation rule for Date of Birth for Adult and Dislocated Worker.	The Department does not agree with this comment universally. In some instances, the Department already allows for documentation to support the validity of certain elements. It appears that the state is including exact date matches in the scope of its comment, which	NY

13	The Department's data element validation policies do not fully align with our state's service integration initiative.	is not appropriate. The proper collection and reporting of key dates (i.e., training start dates, program end dates, and dates individuals became employed) are key to establishing correct cohort information for program exiters for performance reporting purposes. This information should be exact and be supported by specific documentation as prescribed by the handbook. The Department is unable to understand this statement without further context. Data validation should be easier to complete if a state is truly integrated, as programs should have greater access to the necessary data through this integration. Further, a common case management system would ensure that the data elements and values collected for a particular individual were the same across programs.	NY
14	The Department should allow "staff verification" as an acceptable method of data element source documentation. The use of staff verification to validate a data element enhances and streamlines the process in which a data element is validated, as well as safeguards storing copies of documents which contain personal information.	As mentioned earlier, the Department supports efforts to reduce the amount of hard copy documents/paperwork required to effectively administer its programs. However, a system that has no verifiable source documentation defeats the intent of data validation and makes states reports/results unverifiable, which is exactly what would occur with the sole use of staff verification, in place of other digital or paper-based documentation. The purpose of the documentation is to allow independent verification that the information input into the reporting databases is accurate. If staff were allowed to verify their own data entry and there is no documentation collected to support critical data elements it will be impossible to independently verify that the information is accurate.	NY
15	Veteran status should only be subject to data element	This comment actually contains two issues. The first	NY, WI, OR

#	Comment	Response	
	validation as related to specific veteran services provided by Disabled Veteran Outreach Program (DVOP)/Local Veterans Service Employment Representatives (LVERs). DD-214 documentation should not be required if services would have been available to them regardless of veterans status.	relates to the need for data element validation for veteran status for DOL programs other than services provided through the DVOP and/or LVERs. With the passage of the Jobs for Veterans Act, priority of service for veterans applies to all of the Department's programs. Since veteran status can be used to determine which individuals do and do not receive services from any program when there are resource limitations, validation of veteran status is appropriate for all of the Department's programs. The second issue contained in this comment relates to the DD-214 as the documentation required to validate veteran status. The Department recognizes that there are acceptable alternatives to the DD-214 and its Veterans Employment and Training Service (VETS) intends to publish further information about those alternatives.	
16	States have never received any analysis of data element validation results — how other states are doing; common errors/error rates, etc. We have no way to measure our progress or compare results to other states.	The Department has not provided any analysis to date, as it has not yet established error rates. The Department feels it needs two years of data under the common performance measures as baseline data to establish reasonable error rates. As the Department prepares to transition to the integrated reporting formats that are part of the WISPR system, it plans to incorporate data validation standards and acceptable error rates after implementation of the new reporting system.

Although no nationwide analysis has been provided as of yet, states do receive feedback from the Department during on-site data validation monitoring reviews and through phone and e-mail exchanges. States receive all | WI |

		of the information they should need from the software to monitor their own results to make improvements to their data collection and reporting practices, so they will be able to meet established error rates when such rates are put in place.	
17	REVIEW FOCUS 4: Minimize the burden of the collection of information on those who are to respond, including through the use of appropriate automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submissions of responses. The data validation software (DRVS) is difficult to use and not released in a timely fashion. There are many versions released and states are constantly waiting for patches or upgrades to create reports and perform data validation. Also, many user friendly pieces of functionality available in prior versions of the software have disappeared.	The DRVS is updated to reflect any changes to reporting or validation requirements that occur and also tries to incorporate more functionality as requested by the states. As major changes have been made to the reporting specifications to report on outcomes, and then the common performance measures, the software has been in a relatively constant state of change. In addition, the software had to be re-written, as the underlying software components had reached end-of-life and were no longer being supported by their respective vendors. With all of this occurring at the same time, the focus has been to ensure that the core functionality necessary to calculate and validate the outcomes was in place, rather than focusing on user-friendly "nice-to-have" enhancements. As the report specifications and sampling methodology used in the software are complex, the code associated with the software is extremely detailed and complex, which lends itself to errors being made in the code. This	NJ, PA, WI, IL

		results in multiple versions of the software being released, to not only address new reporting and validation requirements, but to fix issues identified with the software and add user-friendly enhancements as time and resources allow.	
18	The Department does not provide enough training or technical assistance.	Since the inception of data validation, the National and Regional offices have provided ongoing training and technical assistance to states and grantees on all data validation programs. The Department provided intensive training during 2003 and 2004 using webinars and regional meetings to train staff on the software and related requirements. During 2005 and 2006, five of the six regions hosted data validation roundtables to allow states the chance to work directly with ETA staff, review error rates for states in each region and exchange best practices among the states.

Beyond this training, the Department provides technical assistance on data validation through a contractor. During the past three years, the contractor and the Department have responded to over 103,500 technical assistance requests from states and grantees about data validation. These requests can come in through phone, e-mail and/or in-person requests during regional and state meetings. There is no charge or limit to the amount of assistance any one state or program can receive and the Department is prepared to continue providing technical assistance as required for data validation. | PA, IL, OR |

	NATIONAL PROGRAM COMMENTS/ISSUES	ETA RESPONSE	GRANTEES PROVIDING COMMENTS
	REVIEW FOCUS 1: Evaluate whether the proposed collection of information is necessary for the proper performance of the functions of the agency, including whether the information will have practical utility.		
19	Data validation is an essential element to any reporting system to ensure consistency, accuracy and quality.	The Department is encouraged by the grantees comments, as none of them took issue with the overall concept and value of data validation, which demonstrates that they understand and appreciate the importance of data validation.	Able, Experience Works!, Senior Service America
20	The proposed collection of additional extensive information is not necessary; validation is part of the grantee's daily, quarterly and annual quality and internal reviews.	Data collection and appropriate documentation in the client files is necessary to operate and manage the program and prove that funds are being spent to serve the intended population. This is not a new requirement imposed as a result of the data validation initiative. By having all SCSEP grantees perform a standardized data validation methodology, the Department is able to support the quality of the information used to assess the effectiveness of the SCSEP program.	Able, Experience Works!
21	The Department should conduct a data validation pilot for the Senior Community Service Employment Program (SCSEP) in the first year to establish a baseline. If data validation is applied retroactively to Program Year (PY) 2006 data using PY 2007 common measures, the information gleaned and the resulting error rate should not be construed	SCSEP grantees will undergo a full-scale pilot testing to help them and their sub grantees master the process and to build the appropriate procedures to conduct validation. Pilot testing will enable the Department to collect valuable grantee input for refining the software, source documentation	Able, National Urban League, Senior Service America, National Indian Council on

as a measure of overall data reliability on the part of national or state grantees.

		Aging

requirements, list of elements to be validated and other aspects of the methodology. Grantees will not be officially accountable for eligibility or performance data captured prior to the distribution of the validation requirements.

REVIEW FOCUS 2: Evaluate the accuracy of the agency's estimate of the burden of the proposed collection of information, including the validity of the methodology and assumptions used.

		Experience Works!

22 The assumption that grantees spend six minutes to collect each element is inaccurate for SCSEP grantees. Without knowing the sample size, determining the required average number of hours to validate data is not possible, but 106 hours as stated in the Federal Register does not seem realistic based on the volume of eligibility and performance data that requires some sort of documentation which will have to be reviewed through the proposed data validation process.

As indicated in Table 3 in Section 12.B of the Supporting Statement, the average annual burden across the three national program grantees (the National Farmworker Jobs Program [NFJP], the Indian and Native American program and SCSEP), is actually 103 hours (not 106). The burden estimate for SCSEP grantees only is estimated at 162 hours, which includes the time for validators to review sampled case files and 15% of a supervisor's time. Since these are averages, the actual amount of time to conduct data validation will vary slightly by grantee and may increase or decrease after the program actually implements data validation.

Furthermore, the commenter appears to be lumping documentation requirements needed to make eligibility determinations together with data validation, when they are not one in the same. Documentation is necessary to have in the client files to prove the individual was eligible for the program/funds are being spent to serve the intended

		population. Data validation is focused on determining the accuracy of performance outcomes; in other words, it only verifies an individual's eligibility for the program to the extent necessary to determine that they have been/should be appropriately included in the SCSEP performance calculations. It should be noted that the commenter is correct in saying that there are a number of eligibility elements that could be validated for each individual, however, in the overwhelming majority of participants' records, only two or three of these elements are actually populated and would need to be validated. Data validation is focused on determining the accuracy of performance outcomes; and as such, the Department will look to minimize the validation of elements around an individual's eligibility for the program to the extent necessary to determine that they have been/should be appropriately included in the SCSEP performance calculations when finalizing the list of data elements to be validated for SCSEP.	
23	The Department should push back the implementation date beyond the end of Calendar Year 2007 to allow for quality training and to allow grantees to develop new systems and reporting based on the extensive changes in metric goal setting.	The Department implemented data validation for state workforce programs back in Program Year (PY) 2003, but delayed implementing validation for SCSEP until their reporting and case management system, SPARQ, was fully operational. As this new system has been rolled out and there are no other significant changes on the horizon, the Department does not understand why there would be a need to further delay the implementation of validation for	Able, Senior Service America, National Indian Council on Aging

#	Comment	Response	Commenter
24	The Department should consider that SCSEP faces unique challenges in the workforce system including large case loads per professional with participants often located remotely from staff. As such, increasing the document gathering burden for data validation purposes will pose a significant barrier, especially for rural programs	SCSEP to develop a new system. As mentioned previously, data collection and appropriate documentation in the client files is necessary to operate and manage the program and prove that funds are being spent to serve the intended population. This is not a new requirement imposed as a result of the data validation initiative. The Department is required under the Government Performance and Results Act to demonstrate convincing results to Congress annually as part of the appropriations process. In order for these results to be considered convincing in nature, agencies are required to ensure the quality, objectivity, and integrity of the data presented to Congress. This includes all programs, like SCSEP, funded and administered by the Department. Further, large caseloads of front-line staff should not be an important consideration as staff should not be conducting the validation of the data they entered. Large caseloads do not mitigate the Department's commitment to data quality.	Able
25	REVIEW FOCUS 3: Enhance the quality, utility, and clarity of the information to be collected. Several grantees made requests that the Department re-examine specific data elements currently proposed to be validated for the SCSEP program. These include: 1) For the "most-in-need" data element, data validation	The Department will take these comments into consideration when finalizing the list of elements to be validated for SCSEP and determining the appropriate source documentation.	National Urban League, National Indian Council on Aging, Senior

#	Comment	Response	Commenter
	should be required in case notes, clearly describing factors for *each* determination, and be signed off by the project director. 2) Some grantees support a notarized statement and letter from the Social Security Administration for all enrollees who claim income eligible via self-attestation. Grantees have experienced several difficult situations during recertification with enrollees who previously declared self-attestation without documentation. 3) Two elements, "Frail" and "Severe Disability", appear to be identical characteristics. Please consider validating only one of these elements to reduce burden.	Grantees have the ability to require more stringent documentation than what is required by the SCSEP data validation requirements. If a grantee wants to require a notarized statement and letter from the Social Security Administration in lieu of self-attestation, that is acceptable. The commenters are incorrect in stating that "Frail" and "Severe Disability" are identical characteristics, they are not. Grantees should refer to the <u>Interim Final Rule, (pub. 6/29/07)</u> which establishes the new SCSEP performance accountability measures mandated by the 2006 Amendments to Title V of the Older Americans Act. All of the definitions for the "most in need" data elements are defined, including frail and severe disability.	Service America
26	REVIEW FOCUS 4: Minimize the burden of the collection of information on those who are to respond, including through the use of appropriate automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submissions of responses. Case notes should be acceptable for data validation where indicated in the current information. Self-attestation should be permitted for the fields indicated, and grantees should be able to design their own form or tool for gathering self-attestation with no further documentation required.	The Department will distribute a SCSEP data validation handbook that contains a more detailed description of the validation methodology, the overall validation process, procedures and security issues, as well as the source documentation requirements for each data element to be validated. The handbook will also include specific guidelines regarding the use of case notes and self-attestation to	Able, Senior Service America, National Indian Council on Aging

		support certain fields. As mentioned before, grantees have the ability to require more stringent documentation than what is required by the SCSEP data validation requirements.	
27	The grantee believes SCSEP and WIA programs should have identical documentation requirements and access to electronic data for specific data elements to minimize the burden on the client.	The Department agrees and, in so far as the definitions for specific data elements are the same, SCSEP will adopt WIA source documentation requirements when finalizing the specific data elements to be validated /appropriate source documentation for SCSEP.	Experience Works!