

National Core Indicators®

2018-19 In-Person Survey (IPS)



Overview

Table of Contents

NCI History and Activities	1
Overview of National Core Indicators.....	1
State Participation.....	2
The Core Indicators.....	2
Domains and Sub-Domains.....	3
How NCI Data Are Used	3
Methodology.....	4
Administration	4
Proxy Respondents	4
Surveyor Training.....	5
Data Analysis.....	5
Criteria for Exclusion of Responses.....	5
Responses for Section I	5
Responses for Section II.....	5
Sampling.....	6
Why do we ask states to collect 400 surveys?.....	6
Data Management and Analysis	8
Weighting.....	8
Significance Testing.....	8
Technical Details	9
Presentation of Data	9
Important Note on Missing Data	9
Outcome Adjustment.....	9
Scale Development	10
Collapsing Data	10

NCI History and Activities

Overview of National Core Indicators

In December 1996, the National Association of State Directors of Developmental Disabilities Services (NASDDDS), in collaboration with the Human Services Research Institute (HSRI), launched the Core Indicators Project (CIP). The aim of the project was to support state developmental disabilities operating agencies in the development and implementation of performance and outcome indicators—and related data collection strategies—so that they could measure service delivery system performance. This effort, now called National Core Indicators (NCI), strives to provide states with valid and reliable tools to help improve system performance and better serve people with intellectual and developmental disabilities and their families. Moreover, NASDDDS' active sponsorship of NCI facilitates pooled knowledge, expertise, and resources among the states.

In 1997, 15 states convened to discuss the scope and content of a potential performance measurement framework. Directors and staff from these 15 states worked to identify the major domains and sub-domains of performance, indicators, measures, and data sources. The original 61 indicators, developed through a consensus process, were intended to provide a system-level “snapshot” of how well each state was performing. The states were guided by a set of criteria that was designed to select indicators that were:

1. Measurable
2. Related to issues the states had some ability to influence
3. Important to all individuals they served, regardless of level of disability or residential setting.

During this initial phase, data collection protocols were developed and field-tested, including a face-to-face In-person Survey¹ (for individuals age 18 and older who were receiving services) and a mail-out Adult Family Survey (for families who have an adult family member living at home). Seven states volunteered to pilot test the indicators. Eight additional states served on the Steering Committee.

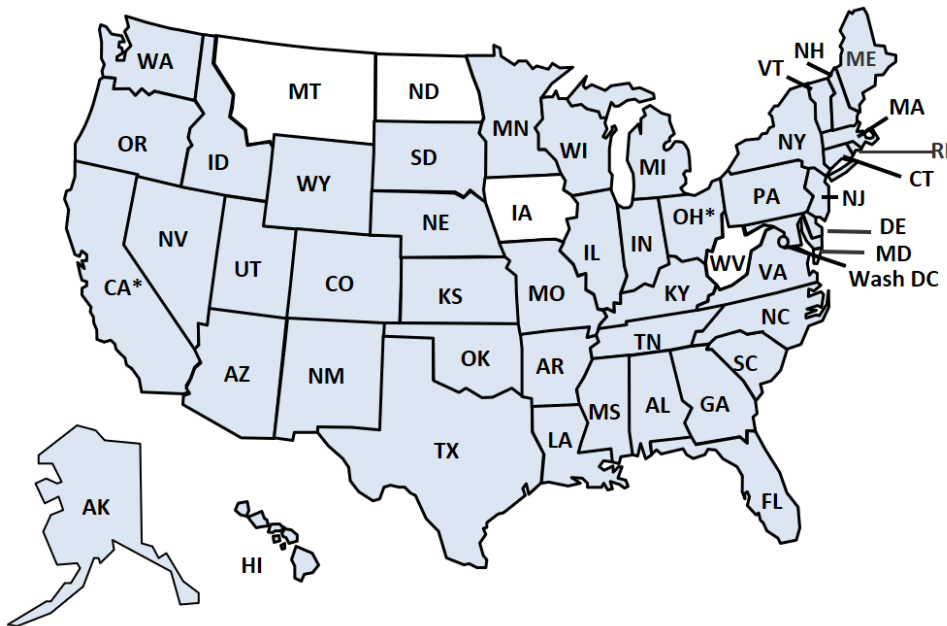
Since the initial field test, NCI has expanded its scope to include outcomes of services for children with intellectual and developmental disabilities and their families. In addition, NCI continues to develop and refine the indicators and expand state participation. For more information about NCI states, technical reports, and other resources, please visit <https://www.nationalcoreindicators.org>.

¹ Formerly named 'Adult Consumer Survey'

State Participation

During the 2018-19 data collection cycle, 46 states, the District of Columbia and 22 sub-state entities participated in NCI. State participation is entirely voluntary, and the participating states are shown in the chart below. NCI states do not participate in all surveys each year.

Figure A. NCI State Participation 2018-19



The Core Indicators

The Core Indicators are the standard measures used across states to assess the outcomes of services provided to individuals and families. Indicators address key areas of concern, including employment, respect/rights, service planning, community inclusion, choice, and health and safety. To see the entire list of Core Indicators, please visit

<https://www.nationalcoreindicators.org/indicators>.

Each survey instrument is designed to measure certain Core Indicators. While most indicators correspond to a single survey question, a few refer to clusters of related questions. For example, the indicator that measures Community Inclusion (the proportion of people who regularly participate in everyday integrated activities in their communities) is measured by several survey questions that ask about several separate community activities.

The current set of performance indicators includes approximately 100 consumer, family, system, and health and safety outcomes—outcomes that are important to understanding the overall health of public developmental disabilities agencies. Indicators are organized across domains: Individual Outcomes; Health, Welfare and Rights; System Performance. Each domain is broken down into sub-domains (see Figure B on the following page). Three data sources are used to assess outcomes: the In-person Survey, three Family Surveys, and a Staff Stability Survey (e.g., staff turnover).

The indicators have remained generally consistent over the last several years and thus can be used to analyze system-level trends over time. However, the NCI program is a dynamic effort that allows for measures to be added, dropped, or changed to reflect current and future priorities of participating states.

The data collection tools used to gather indicator data are regularly refined and tested to ensure they remain valid, reliable, and applicable to current issues within the field.

Domains and Sub-Domains

The following table lists the domains and sub-domains covered by the NCI In-person Survey indicators.

Figure B. In-person Survey Indicators: Domains and Sub-Domains

Individual Outcomes Domain

Sub-domain	Concern Statement
Work	People have paid jobs in community-based settings or have otherwise meaningful day activities.
Community Inclusion, Participation and Leisure	People participate in activities in their community and have opportunities to do things that they enjoy in the community.
Choice and Decision-Making	People make choices about their lives and are actively engaged in planning their services and supports.
Self-Direction	People participate in directing their own supports and services.
Relationships	People have friends and relationships and are able to maintain their friendships and relationships.
Satisfaction	People are satisfied with their everyday lives – where they live, work, and what they do during the day.

System Coordination

Sub-domain	Concern Statement
Service Coordination	Service coordinators are accessible and responsive to people. The service plan is responsive to people's goals and needs. People participate in the service planning process.
Access	Services and supports of quality are readily available.

Health, Welfare, and Rights

Sub-domain	Concern Statement
Safety	People feel safe.
Health	People secure recommended health services.
Medications	Medications are used effectively and appropriately.
Wellness	People maintain healthy habits.
Respect/Rights	People receive the same respect and protections as others in the community.

How NCI Data Are Used

The Core Indicators provide information for quality management and are intended to be used in conjunction with other state data sources, such as risk management information, regional level performance data, results of provider monitoring processes, and administrative information gathered at the individual service coordination level. States typically use the indicator data to inform strategic planning, produce legislative reports, and prioritize quality improvement initiatives. Some states use NCI as a data source for supplemental performance measures in their home and community-based services (HCBS) waiver quality management systems and include the information in support of evidentiary reports to the Centers for Medicare & Medicaid Services (CMS). Many states share the indicator data with stakeholder groups such as Quality Councils and use the stakeholder feedback to help set priorities and establish policy direction. It is also important to note that states do not use the information in a punitive way to sanction service

providers, nor do they use the results to remediate individual issues (unless specifically requested by the participant or required by law as in the case of suspected abuse, neglect, or mistreatment).

For more information on how to use these data for quality improvement, please see this handbook: [Using National Core Indicators for Quality Improvement Initiatives](#).²

Methodology

Administration

A key principle of NCI is the importance of gathering information directly from service recipients. The indicators are meant to apply to all individuals receiving services from the state developmental disabilities service system, regardless of disability type, and NCI administration protocols dictate that every person selected into the sample is given an opportunity to respond and no one is prescreened or predetermined to be unable to respond. Moreover, survey questions should not be marked “not applicable” on the basis of a person’s level of ability.

Information for the In-person Survey is collected via a direct conversation with the person receiving services. Additionally, background information is primarily collected from the individual’s record. Section I questions, which pertain to personal experiences and require subjective responses, may only be answered by the individual receiving services. Section II of the survey—which consists of objective questions on the individual’s involvement in the community, their choices, respect and rights, and their access to services—allows for the use of “proxy” or other respondents who know the individual well (such as a family member or friend).

Proxy Respondents

Studies have found that the greatest discrepancies between individual and proxy responses occur when the information being collected is subjective (i.e., related to how a person feels; proxies would be aware of the correct answers only if the individual had previously expressed his or her feelings).³ Questions relating to observable behaviors tend to have higher levels of agreement between the proxy and the individual. By excluding proxies, a large percentage of individuals (most of whom are unable to respond) would be unrepresented in the data. Thus, the NCI project determined at the outset that proxy respondents would be used, though only for specific questions and only in situations where the individual surveyed either could not effectively communicate with the surveyor or chose to have a proxy respondent.

As mentioned, the use of proxy respondents for the NCI tool is limited to questions in Section II, which relates to observable and/or measurable items: Community Inclusion, Choices, Respect/Rights, and Access to Needed Services. These questions deal with everyday occurrences on which others may reliably report. State records are also used to report objective data on an individual’s health status and exam history as well as employment status; this information is collected in the Background Section of the survey.

NCI has taken steps to ensure that the responses gathered are as accurate as possible. First, it is important to make the questions as accessible as possible to increase participation by all individuals. NCI attempts to make the language in the In-person Survey as easy to understand as possible and includes suggested rephrasing for questions that may be nuanced or more difficult to understand. The NCI surveys are also routinely revised based on feedback from states, self-advocates, families, and surveyors who are administering the tool. Secondly, it is important to establish a set of standards for proxy respondents. NCI aims to increase the reliability of proxy responses by allowing only those people who know the individual well (family, friends, staff, etc.) to serve as proxy respondents.

To avoid conflict, service coordinators are not allowed to respond as proxies. Further, if both the individual and a proxy respondent answer a question, the individual’s answer is recorded so long as his/her answer is deemed reliable by the

² Located on the National Core Indicators website: www.nationalcoreindicators.org → Resources → Technical Reports

³ Magaziner, Jay, Sheryl Zimmerman, Ann L. Gruber-Baldini, J. Richard Hebel, and Kathleen M. Fox. "Proxy Reporting in Five Areas of Functional Status Comparison with Self-Reports and Observations of Performance." *American Journal of Epidemiology* 146.5 (1997): 418-28.

surveyor. Surveyors also keep track of who responds to each question—the individual or the proxy—and the proxy’s relationship to the individual.

Surveyor Training

States employ a variety of individuals to conduct the face-to-face meetings for the NCI In-person Survey. To avoid conflict, surveys cannot be administered by the individual’s service provider, relative, personal case manager, staff member, or other close contact. Given this constraint, states often use university students, state staff, private contractors, advocacy organizations, and individuals with disabilities and their families to conduct the surveys.

HSRI and NASDDDS staff provide two sets of trainings each year: a more comprehensive training for new states and states with mostly or all new surveyors and a web-based refresher training for surveyors who have recently participated in the In-person Survey. Typically, instruction takes the form of a “train-the-trainer” session, whereby staff from the state’s lead agency and/or those managing the surveyors provide training to new hires throughout the year. The goal of trainings is to provide a standardized training to ensure uniform application of the survey.

Trainings consist of a detailed review of the survey tool and an overview of general surveying techniques. Trainers are given a packet of standardized materials to be used at the on-site training sessions. An additional session reviews the process for entering survey data in the Online Data Entry Survey Application (ODESA).

Data Analysis

Criteria for Exclusion of Responses

All individuals selected in the survey sample are given an opportunity to participate in a face-to-face meeting. There are no prescreening procedures. Exclusion of responses occurs at the time of data analysis by HSRI, based on the criteria described below. There is no threshold of number of answers to be given in order to consider a survey complete.

Surveys are excluded from analysis of questions occurring in Section I if:

1. The surveyor indicated that the individual receiving supports did not respond validly to questions in Section I.
2. All questions in Section I were missing or marked n/a or Don’t Know.

Surveys are excluded from analysis of questions occurring in Section II if:

1. The individual receiving supports was marked as the respondent to all questions in Section II but Section I was deemed invalid (for one of the reasons above).
2. No questions were answered in Section II.

The total number of valid In-person Surveys administered in 2018-19 was 22,009.

Responses for Section I

Only individuals receiving services may answer questions in Section I. As shown in Figure C, 71.6% of individuals in the total 2018-19 sample were able to respond to Section I of the direct meeting; valid responses by state ranged from 54.2% to 96.2%.

Responses for Section II

Section II allows for multiple respondents who know the individual well (e.g., family, friend, support worker) to provide answers. If an individual’s responses were excluded from Section I in the final analysis, responses from Section II were also excluded if the individual was the only respondent to Section II. For 2018-19, the valid response rate (proxies included) to Section II was 98.2%.

Sampling

Each state is instructed to attempt to complete a minimum of 400 surveys with a random sample of individuals over the age of 18 who are receiving at least one publicly funded service besides case management. A sample size of 400 guarantees valid comparisons to be made across states with a 95% confidence level and a margin of error of +/- 5%, no matter how large the service population size. A 95% confidence level and a margin of error of +/-5% is mandatory for each state's sample to be included in this report. Most states sample more than 400 individuals to account for refusals and surveys that may be deemed invalid. Figure C presents the number of surveys completed by state. A more detailed description of each state's sampling strategy can be found in Appendix B.

Why do we ask states to collect 400 surveys?

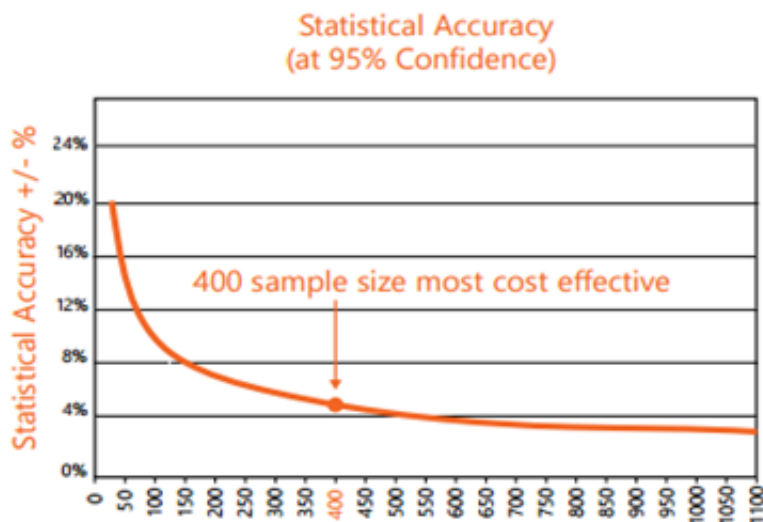
When NCI asks states to pull a representative sample of 400 surveys, the aim is to compile a dataset that can provide a high level of statistical accuracy. This means that the goal is to create a dataset that can be used to make accurate inferences about the entire population of people using public services in each state. In order to do this, we ask each participating state to pull a sample that reaches the 95% confidence level/5% margin of error threshold; for most states, this sample size is 400.

A sample size that reaches the 95% confidence level/5% margin of error threshold means that if you sampled that number of people from the same population several times, 95% of your responses would fall within 5% above or below the actual percentages. Using this sample size we can assume relative accuracy. The larger the sample size, the smaller the margin of error and the greater the accuracy.

Statistical accuracy is largely a function of sample size. The larger the sample size, the greater the statistical accuracy of results (lower margin of error). However, gains in statistical accuracy are not proportional to increases in sample size. There are diminishing returns. For example, for a large total population, if the sample size increases by 100, and then by another 100, and then by another 100, the statistical accuracy might increase by 5%, then by 3%, then by 1%.

As a result, the 400 sample size gives a statistical accuracy of $\pm 5\%$ and is considered the most "cost effective" sample size for large-scale surveys. Meaning, you could sample more people, but the increase in statistical accuracy those "extra" people would provide is minimal, and not worth the money.

To play around with this, take a look at www.raosoft.com/samplesize.html NCI requires a margin of error at most 5% and a 95% confidence level. As you increase the population size (in our case, this represents the service population), you'll see that the recommended sample size does not increase proportionally.



<https://answersresearch.com/wp-content/uploads/2015/06/Answers-Research-Magic-400-062315.pdf>

Figure C. Valid Surveys and Response Rates by State 2018-19

State	% Valid Answers to Section I	% Valid Answers to Section II	Number of Valid Surveys Completed
AL	91.0%	99.7%	378
AR	78.4%	99.4%	510
AZ	71.2%	100.0%	413
CO	95.3%	97.9%	712
CT	71.4%	96.6%	391
DE	71.1%	95.6%	342
FL	70.9%	98.5%	867
GA	68.9%	99.6%	479
HI	45.2%	99.4%	363
IN	71.2%	99.1%	742
KS	68.5%	98.8%	408
KY	62.7%	100.0%	429
ME	83.8%	95.0%	400
MI	63.1%	99.7%	658
MN	73.9%	94.8%	2140
MO	72.2%	100.0%	403
NC	57.1%	99.8%	651
NE	74.4%	99.5%	418
NH	77.0%	98.0%	400
NJ	65.6%	99.6%	514
NV	84.8%	98.3%	408
NY	54.2%	94.4%	448
OH	74.2%	99.0%	732
OK	69.0%	99.5%	400
OR	78.9%	99.8%	421
PA	73.8%	98.9%	710
RI	80.7%	98.9%	358
SC	96.2%	100.0%	501
SD	88.0%	93.4%	351
TN	81.8%	100.0%	406
TX	59.8%	100.0%	2388
UT	63.4%	98.7%	372
VA	73.9%	99.8%	807
VT	74.3%	98.6%	346
WA	72.1%	99.3%	441
WI	67.2%	94.4%	961
WY	80.4%	92.4%	341
Unweighted Total	71.6%	98.2%	22,009

Data Management and Analysis

HSRI coordinates the NCI data management and analysis. All states enter data into the Online Data Entry Survey Application (ODESA) system; HSRI staff subsequently download the data into an SPSS data file. This data file is then “cleaned” (reviewed for completeness and compliance with standard NCI formats) and invalid responses are eliminated. Files from individual states are then merged into a single SPSS file. The merged file from 2018-19 was used for all analyses in this report.

Below is a summary of the statistical procedures used to analyze the In-person Survey data.

Weighting

Statistically, the term “average” refers to a calculated central or middle value of a set of numbers. In NCI reports, we use “NCI average” to demonstrate the typical performance of all the states that conducted the survey. Prior to the 2016-17 survey cycle, the NCI average was calculated as the simple arithmetic mean of all state means (an approach known as “average of averages”).

The NCI averages are “weighted” to reflect the states’ relative population and sample sizes. We created the weights using the state’s number of valid surveys and its total survey-eligible population. This way, a state that provides services to a larger number of people but uses a sample similar in size to other states has a greater influence on the overall NCI average (that is, its contribution is *proportional to its service population*).

When a state’s sampling strategy is to identify and interview survey participants using simple random sampling or proportional stratified random sampling, each completed survey in the state gets assigned the same weight. In some cases, when a state’s sampling strategy departs markedly from simple random or proportional, it may be necessary for completed surveys in the state to be assigned different weights based on which sampling strata they correspond to.

This survey year, IN, MN, TX and WI intentionally oversampled one or more subpopulations or geographical regions to get more focused insights, so their weights are calculated slightly differently (using the number of completed surveys and service population sizes in each subpopulation or region). You can read more about each state’s sample in Appendix B.

The number of respondents (N) for each question is not weighted.

Significance Testing

For most items in this report, each state’s percentage was compared to the weighted NCI Average (described above), and the differences between the two were tested for both statistical significance as well as effect sizes. Effect sizes are used in addition to statistical significance because statistical significance of a state’s result depends, in part, on the size of the state’s sample: the larger the sample, the more likely it is that even a small difference will be found statistically significant. A statistically significant difference for a state with a large sample size, in and by itself, does not necessarily mean there is a practically significant difference. The inclusion of effect sizes as a new criterion allows us to present “meaningfully significant” results, which take into consideration the magnitude or size of the differences.

Many outcomes in this report are categorized into one of the three classes:

1. Significantly above the NCI average, where the difference between the state’s percentage and the weighted NCI average a) was in favor of the state, b) was statistically significant (i.e., $p < .01$), and c) met the effect size criterion (i.e., Cohen’s $d > 0.2$, see below for details);
2. Within the NCI average range, where the difference between the state’s percentage and the weighted NCI average was a) not statistically significant (i.e., $p \geq .01$), or b) did not meet the effect size criterion (i.e., Cohen’s $d \leq 0.2$);
3. Significantly below the NCI average, where the difference between the state’s percentage and the weighted NCI average a) was in favor of the NCI average, b) was statistically significant (i.e., $p < .01$), and c) met the effect size criterion (i.e., Cohen’s $d > 0.2$).

Technical Details

The comparisons were done through one sample T-tests using the weighted NCI average as the benchmark. A conservative cut-off point (alpha) of $p < .01$ was used to detect statistically significant differences. Effect sizes are calculated using the formula: $\text{Cohen's } d = (2t) / (\sqrt{df})$. A cutoff point of Cohen's $d = 0.2$ was chosen for the effect size to be considered "meaningfully significant," following the standard interpretation offered by Cohen (1988) that Cohen's d of $0.8 =$ large, $0.5 =$ moderate, and $0.2 =$ at least a small effect.

Presentation of Data

Tables that show significance also break out average results for each state by the type of residence: ICF/IID or other institutional settings, group residence, own home, parent/relative's home, foster or host home. Figure A on the following page shows, by state, the percentage of people surveyed in each residence type. Please note though that:

1. Significance testing was not performed on the breakouts by residence
2. The number of people in each residential setting is often too small to allow for valid state-to-state comparisons
3. And in many cases, statistically valid conclusions cannot be drawn about differences in results between residence types

Still, we provide this information for states' internal purposes. It should not be used to compare one state with another.

For outcomes that were not tested for significance, tables show state outcomes listed in alphabetical order.

To protect people's privacy, we do not show the data when fewer than 20 people from a state or residence type responded to a question (though their responses are included in the NCI average).

Important Note on Missing Data

For items in this report that come from administrative records, several states had large amounts of missing data or data recorded as "don't know." Where 25% or more of a state's sample had "don't know" or missing responses, we used an asterisk (*) to indicate that state's data should be interpreted with caution as it may not accurately represent the sample or service population.

Due to the high rate of missing data and "don't know" responses in the Health, Wellness, Medication, and Employment sections, some tables are shown twice. The first table illustrates data as typically reported (by significance tier and averages in descending order, excluding missing data and "don't know" responses). The second table shows data in alphabetical order and includes the percentage of missing and "don't know" responses. Data in both sets of tables are weighted.

Outcome Adjustment

Outcome adjustment or "risk adjustment" is a statistical process used to control for differences in the individual characteristics of people who completed the NCI Survey across states. This procedure allows for more legitimate state comparisons by effectively "leveling the playing field." This analysis helps account for the fact that states have different eligibility definitions for services and may have samples with different characteristics. Only those indicators that are likely to be affected by individual characteristics were adjusted (e.g., indicators in the Choice sub-domain). For example, a person who has limited mobility and is older may be less likely to choose where they live. On the other hand, such characteristics should not affect whether a person has friends or staff who are respectful.

The indicators are risk-adjusted using the following characteristics: age, level of intellectual disability, level of mobility, and whether any behavioral supports are needed to prevent self-injury, disruptive, or destructive behavior. Risk adjustment was only done on Choice items for this report. Results for indicators that were adjusted are labeled as such, and unadjusted results for these indicators are presented in Appendix C.

Outcome adjustment produces a predicted value that one would expect to observe given the individual's characteristics. The state's average observed rate (i.e., the state average prior to risk-adjustment) was adjusted by the average

predicted rate to produce the risk-adjusted rate. As a result of this procedure, state differences in adjusted values reflect true state differences rather than differences due to the demographic or characteristic makeup of state service populations.

Scale Development

Scales are used to combine responses from multiple similar questions into one variable. In this way we can measure an overarching concept rather than specific elements of the concept. For the In-Person Survey, scales are used for the Choice and Decision-Making sub-domain. To create a scale, statistical tests are required to ensure that the multiple items share common properties. Scales were created using factor analysis.

The Choice and Decision-Making sub-domain includes two scales: Life Decisions and Everyday Choices; The Choice scales were created by averaging the items within each category. Additionally, one Community Inclusion scale was produced by adding four of the items.

Higher scale scores represent higher levels of choice. The scales were also risk-adjusted. Each scale's reliability was assessed using Cronbach's alpha. The scales had an adequate level of internal consistency and reliability.

Collapsing Data

For many of the indicators in this report, only "yes" responses are analyzed and reported for state to state comparisons and significance testing. However, there are several indicators for which the "yes" response was collapsed with the middle response (for example, a "sometimes" response, or a "some input" response) to form the "yes" indicator category which was then analyzed and subsequently reported. For example, for choice indicators, responses "person chose" and "person had some input" were collapsed into one "yes, person chose" category. For those indicators, the two responses were considered to be equally indicative of a positive outcome. Appendix A contains the collapsing and recoding rules for all indicators and risk-adjustment variables. Appendix C shows all un-collapsed data.