

# Focusing on Outcomes in Centers for Independent Living:

Report from the National Council  
on Independent Living  
*Outcome Measures Task Force*



April 6, 2012



Since 2006, the National Council on Independent Living (NCIL) has sponsored a concerted effort to measure and improve program outcomes in the approximately 400 Centers for Independent Living throughout the United States. Four factors have motivated this emphasis on outcomes:

- Centers themselves were calling for a better way to capture their accomplishments, not only to share their successes with various audiences, but also to improve their own effectiveness.
- The Rehabilitation Services Administration (RSA) of the U.S. Department of Education (ED) had begun an increased focus on outcomes, but not all Centers agreed with the desired outcomes and the measurement methods RSA was suggesting.
- In 2003, the federal Office of Management and Budget (OMB) had applied its Program Assessment Review Tool (PART) to the Center program within RSA, and OMB concluded that the program suffered from “Results not demonstrated”. This OMB finding was explicitly *not* a statement that the Center program was being ineffective, but simply that the program was not yet identifying its desired outcomes, measuring them, and using that information to continually improve effectiveness.
- Some states and Centers had begun their own, early efforts to focus more on outcomes, and those results were promising enough to encourage a more standardized, national effort.

### **Initial Organizing and Planning**

National Council on Independent Living Task Force. Recognizing the importance of this issue, NCIL established an Outcome Measures Task Force, chaired by a highly experienced and respected veteran of the Independent Living (IL) field, and comprised of 10 nationally known representatives from Centers, Statewide Independent Living Councils (SILCs), support organizations such as NCIL and the Independent Living Research Utilization (ILRU) a program of Memorial Hermann, and academia. The Task Force was funded by NCIL, ILNet, and the University of Kansas RTC/IL, and it recruited a national expert in program outcomes as its independent consultant.

For its earliest, in-person meetings, the Task Force also invited representatives from both RSA and OMB. Both organizations accepted this invitation, and their representatives attended and added a great deal to the discussions.

Philosophically, the Task Force agreed from the beginning that Centers should aim to practice outcomes management, not just outcomes measurement. The difference is between a research activity that simply measures outcomes and reports them to various audiences (outcomes measurement) and an organizational development activity, integrally intertwined into Center operations, that not only measures progress on its desired outcomes but also uses that information to identify program weaknesses, identify possible improvements, choose which improvements are most promising, implement those improvements, then measure outcomes again (outcomes management). Some call this Managing for Results, others call it Results-Based Management, but the intent in all cases is the same: programs should manage and improve outcomes, not just measure them.

Desired outcomes and a logic model of the Center program. Operationally, the Task Force first identified the desired outcomes for the Center program, organized into a logic model. Logic models are enormously popular and useful tools to show visually, on one page, what a program does (its activities), what the program is trying to achieve by conducting those activities (the program’s outcomes), and in

what sequential order the program expects each outcome to be achieved (the different levels and “streams” of outcomes). (Note: some logic models also display program inputs and outputs, but activities and outcomes are the two essential components.)

Since there is no actual, “national” Center program (each Center is tailored to the needs and capacities of its community), the Task Force identified a set of outcomes that could reflect reality as much as possible for as many Centers as possible. The evaluation consultant reviewed the IL literature and created a first draft of a logic model, which the Task Force revised extensively. Since the goal was to include the wider IL community as much as possible, the Task Force sent this next draft to the IL field and encouraged readers to suggest revisions. In addition, the Chair of the Task Force presented and discussed the logic model and its desired outcomes at several national, regional, and state conferences.

After reviewing the many thoughtful comments, the Task Force revised its draft into the current logic model -- one version a visual graphic and another version a text-only description of this visual graphic. To summarize these two versions, our Center program logic model illustrates that the three activity streams of *IL Services*, *Information and Referral*, and *Systems Advocacy* achieve six different initial outcomes, that these initial outcomes lead to achieving seven intermediate outcomes, and that these intermediate outcomes lead to achieving three ultimate outcomes. That is, the lower 15 outcomes converge to achieve the highest-level ultimate outcome: **“Persons with disabilities are integrated into American society”**.

At this point in the project, the challenge changed from defining desired outcomes and their interrelationships to specifying how those outcomes might be measured. To best respond to this different challenge, NCIL revised the Task Force to include experts on measurement within the IL field. Other than the Chair and the evaluation consultant, many of the 12 other members were new.

Choosing key outcomes. The second Task Force began its work by choosing which of the 16 outcomes on the logic model should receive attention first. Obviously all 16 are important, but a program rarely has enough resources to measure progress on each one of its desired outcomes. Instead, a program typically needs to balance which outcomes are most central to its essence, which outcomes will produce information most useful for improving program performance, and which outcomes are of most interest to various audiences. Inevitably, this means that some key outcomes will be chosen, while others will be left for later attention.

Once again the wider IL community was consulted, and via email and the Internet interested persons chose their top six preferences. As sometimes happens with a democratic process, the IL field identified eight outcomes of great interest, not just six. The box at the right shows the eight outcomes chosen.

Measurable indicators. Outcomes are not measured directly, so the Task Force developed one or more measurable indicators for each outcome. Developing good indicators is the heart of any outcomes effort, and four separate approaches were used to develop the best indicators:

- The combined expertise of the Task Force's members and its evaluation consultant
- A fairly extensive literature review of relevant materials
- Personal contact with respected IL researchers in the United States and other countries
- Perhaps most innovative and involving for the wider IL community, a public competition with money prizes for the individuals offering the best suggestions. Dozens of suggestions were received, and the three winners received \$100 each.

As a result of this multi-pronged approach, the Task Force adopted *12 measurable indicators* for these *eight key outcomes*.

Sources and methods for gathering outcome information. The Task Force decided to measure these 12 indicators in five separate ways: (1) telephone interview a random sample of Center consumers (persons with a Consumer Service Record (CSR), or a family member if necessary), (2) telephone interview a random sample of I&R callers (excluding any person with a CSR), (3) review Center documents and files, (4) download data from Center Management Information Systems (MIS's), and (5) interview Center Directors.

Three separate instruments were developed to gather the needed information, including an interview guide for calls to CIL consumers (with four closed-ended and seven open-ended questions), an interview guide for calls to I&R callers (with two closed-ended questions), and an information-gathering form for each Center (with one closed-ended and six open-ended questions). Each of these paper forms was re-created exactly on Survey Monkey, a widely accepted Internet-based survey research program, so a Center could transfer the information from its paper instruments directly into an electronic format.

### **Eight CIL Outcomes Chosen for First Attention**

#### IL Services stream

- Persons with disabilities have skills/knowledge/resources to support their choices
- Persons with disabilities are more independent

#### I&R stream

- Persons with disabilities get the information they need
- Persons with disabilities advocate for increased community supports

#### Systems Advocacy stream

- Barriers, problems identified
- A consumer agenda for change exists
- Decision-makers act on our agenda
- Methods and practices promote independence

## The First Field Test

With any new effort, especially one as complex as focusing on program outcomes, it is essential to conduct a field test of all materials and procedures. Many plans that seem reasonable during discussions or on paper can turn out to have serious problem when tried in reality. Finding those problems sooner rather than later, and certainly before a full-scale effort is launched, is the purpose of a field test.

With that in mind, a five-member subgroup of the Task Force identified 26 real-world challenges that could be foreseen. These included which Center staff or volunteers should gather outcome information, how best to encourage accurate data, how many consumers and I&R callers to interview, how soon after the initial service contact to interview a consumer, what system to use to store outcome information, how best to provide on-going technical assistance during the field test, how to define key terms such as “systems advocacy” and “at risk”, etc. This subgroup recommended actions to the full Task Force, the Task Force accepted these recommendations, and the field test began.

Recruiting Centers to participate. The Task Force made a conscious decision *not* to limit participation in the field test to only NCIL members; all Centers, regardless of affiliation, were invited to participate. The Chair of the Task Force was in charge of recruiting Centers, and he deliberately sought Centers that varied on several dimensions:

- Number of consumers served (350 or fewer vs. 351 or more)
- Number of full-time equivalent (FTE) staff (9 or fewer vs. 10 or more)
- Annual funding level (\$300,000 or less; \$300,001-\$500,000; \$500,001-\$1,000,000; over \$1,000,000)
- Geographic location (urban; rural; both urban and rural)

In yet another demonstration that the Center community recognizes the importance of outcomes, 32 Centers volunteered to participate, even though they received no tangible incentive or reward. Furthermore, each Center agreed to assign two, high-quality staff to the field test; measure all 12 indicators developed by the Task Force; participate in all training and technical assistance activities; use the agreed-upon definitions, procedures, and data management systems; and stay involved to the end of the field test (at least one year).

Training participating Centers. To ensure that everyone gathering outcome information used the same procedures, the Task Force implemented a five-part training plan:

1. An 18-page Training Manual that explained the concepts and procedures to be followed. This manual defined key terms such as “consumer”, “activities to identify or confirm the primary barrier/problems in the community, a “systems advocacy workplan”, and perhaps most problematically “at risk of moving into an institution”. It also explained several specific tasks, such as random sampling, telephone interviewing, calculating percentages, and entering data into Survey Monkey.
2. Two separate, 90-minute conference calls to review the Training Manual and answer any questions before gathering information.

3. Three separate, 90-minute technical assistance (TA) conference calls during the field test for participating Centers to discuss progress and problems, ask questions, and get answers as they gathered information.
4. A private, closed-access listserv for those who prefer that mode of communication and for times between the TA conference calls.
5. The Chair of the Task Force's private email address and phone number for persons who preferred to contact him privately and directly.

Gathering outcome information. Centers were asked to gather outcome information from October 1 to December 31, 2009. However, many Centers did not finish until February 2010 (see Lessons learned below.) Telephone interviewing proceeded with four ethical principles, as noted in the Training Manual:

- *“Always respect the person you’re interviewing, including his or her right to respond in any way at all, even choosing not to participate. The principle of informed consent means that the person you’re interviewing must understand what you’re asking her/him to do, know how his/her information will be used, realize that s/he has the right to refuse to participate without suffering any harm, and freely give her/his permission to proceed.*
- *Treat the person with respect during the interview. There are no right or wrong answers – however the person answers is automatically correct, by definition.*
- *Don’t skew the person’s answers in any direction. We don’t want flattering answers, we want honest answers. Otherwise, we’re wasting their time and ours, too.*
- *Keep confidential all information you learn from personal files and from interviews. Don’t let others in the office hear or see what you’re learning during an interview.”*

Center staff used the paper instruments to record information as they spoke on the phone, then later transferred information from paper into Survey Monkey. Alternatively, Center staff could have entered information directly into Survey Monkey as they conducted the interviews, but a survey research expert advised against that approach.

Analyzing the outcome information. A specialist data analyst worked with the independent evaluator on analyzing the information. Data files were transferred from Survey Monkey to Excel, which allowed more flexibility to conduct more sophisticated analyses. While the data specialist was analyzing the quantitative data, the Chair of the Task Force and the independent evaluator content-analyzed the open-ended, qualitative data from consumers and Centers.

As a result of these analyses, the Task Force produced four separate reports:

1. What types and amounts of information each Center provided to the Survey Monkey database. This analysis both verified the Centers’ work and highlighted any discrepancies between what information a Center thought it had entered and what the computer actually received. When a discrepancy was found, that Center was contacted and the correct numbers calculated.
2. Demographic information about the consumers interviewed overall. To ensure that the consumers interviewed during the field test were representative of consumers nationally, the consumers’ age, race/ethnicity, and disability were analyzed. Knowledgeable Task Force members judged the demographics of the sample to be comparable to overall consumers.

3. Outcome information for each of the 12 indicators. This was the first set of systematically collected outcome data designed by the IL field itself, and naturally it generated considerable interest. These findings (from a more recent field test) are presented later in this report.
4. A tailored, confidential report to each Center, showing how that Center's outcome performance compared to the average performance of all 22 Centers. Progress on each of the 12 indicators was color-coded: green for above-average progress, white for typical progress, red for below-average progress. Centers found this tailored report to be extremely helpful.

The Task Force also analyzed whether certain factors seem to produce better outcomes. For example, are outcomes better for men vs. women? Do younger consumers achieve better outcomes than older consumers? Do consumers with certain disabilities achieve better outcomes than consumers with other disabilities? Despite multiple analyses, no factors were related to better outcomes.

### **Lessons from the First Field Test**

When Centers were recruited for the field test, it was stressed that they had two, quite different roles. As the Training Manual stated, *"In your first role as outcome data collector, you're the person who will actually be collecting the outcome data.... In other words, you're the source of all our outcome data."*

But Centers also had a second role, and again in the words of the Training Manual:

*"You have a second role, though, as field test evaluator, and this role is just as important as your first role. Remember we said that we need to find out if our plans work as well as we hope? Well, you and your colleagues at the 31 other Centers are the only persons who will know the answers to that question.... In other words, only you can conduct this field test and evaluate our current plans. Whether or not we need to change anything, and what we need to change, only you can tell us."*

Debriefing participating Centers. This second role was very important, and the Task Force developed several ways to learn from participating Centers:

- Listening carefully to questions, problems, and suggestions raised on technical assistance phone calls and on the private listserv.
- Asking each Center staff who had conducted telephone interviews to edit each of the three information-gathering instruments into whatever versions they would have preferred. Some staff provided Track Changes to the instruments and others emailed comments. Together they provided many suggestions for changing the wording of questions to the consumers and I&R callers being interviewed, and also the wording of instructions to the interviewers.
- Creating a 23-question Survey Monkey debriefing survey for each Center to complete, including how accurate they think the outcome information they gathered was, how easy or difficult it was to gather the information, whether the process was valuable to their own Center, how the process

could be improved, etc. Each of the closed-ended and open-ended responses to this survey was analyzed, and the findings were provided to each Center.

- Arranging a 90-minute debriefing phone call to discuss the results of this debriefing survey and anything else the Centers wanted to say about the field test.

Two main messages. Centers that participated in the first field test had two main messages for the Task Force. First, and very importantly, **it is possible to measure the outcomes of Center programs**, and the benefits are worth the effort. Sixty percent (60%) of respondents said that participating in the field test had been valuable or very valuable, and 72% were interested or very interested in participating in another field test. Significantly, some Centers were already starting to change their practices, based on what they learned about their own outcome performance.

The Centers' second message, however, was that **focusing on Center outcomes is not simple**. Not all Centers were ready – 10 Centers dropped out before information-gathering began, and one additional Center failed to gather all the information needed. That is, only 21 of 32 Centers (66%) volunteering for the field test were able to provide a full set of outcome information.

Even for these 21 successful Centers, the materials and procedures needed to be improved. Centers took the field test seriously, and they offered many excellent suggestions for improving the entire effort. Below are the major changes made as a result of their guidance.

### **Improvements Based on the First Field Test**

Based on the positive reactions to the first year, the Task Force extended the field test into a second year. However, several improvements were made. In particular, the Task Force:

- Dropped one of the 12 indicators, because it was not interpretable. The indicator read “# and % of open ILP goals achieved within the past federal fiscal year by consumers served by the Center within the first nine (9) months of the past federal fiscal year”. Centers were able to provide this information without difficulty. However, because consumers vary considerably on how many ILP goals they have, it was impossible to calculate a meaningful average number or percentage of goals achieved. For example, some consumers had zero ILP goals, while others had 18 goals. Not surprisingly, then, some consumers achieved zero ILP goals, while others achieved 15. Given this variability, what does it mean that the “average” consumer achieved 1.7 out of 2.7 ILP goals, for an average success rate of 63%? This is an important area, but one that is difficult to measure meaningfully.
- Began to require only one example, not two, for another indicator. Because the Task Force wanted to set a conservative, cautious standard for declaring success on an indicator, it required consumers who essentially said “Yes, I have gained new skills, knowledge, or resources by working with the Center” to document that gain by listing two specific skills, types of knowledge, or resources. Only by listing two (2) separate gains could a consumer receive a positive score on this indicator.



How, though, to count those consumers who answered “Yes, I have made this gain” but could list only *one* specific skill, knowledge, or resource? In order to be conservative in documenting success, instructions required these consumers to be considered a “No”. For Year Two, this requirement was loosened, and consumers were required to list only one specific skill, knowledge, or resource.

- Re-worded many of the questions on the three data-gathering instruments. During the Center debriefing, Centers were clear that the interview questions could be improved. As a result, the wording of many questions was revised to be more in line with specific suggestions from Centers.
- Lowered the number of consumers and I&R callers to be interviewed from 30 to 25 of each, a 17% reduction. Statisticians consulted about this change did not feel the difference would affect the planned statistical tests, so the Task Force reduced the reporting burden on Centers.
- Changed the timing of information-gathering. During the first year, Centers conducted telephone interviews during October-December. While this made sense for substantive reasons, this also gave Centers an extra task exactly when they were preparing annual performance reports to RSA. The Centers asked for a change, and data collection was moved to January-February.
- Improved procedural aspects of telephone interviewing. Originally, Centers looked up background information about a consumer *before* making the phone call; this information was then unused if the consumer could not be located or refused to participate. For Year Two, Centers looked up this information only *after* completing a successful phone interview. Also, up to four Center staff could make calls (twice the number from Year One), and Centers called a particular phone number only three times before moving on to another person.
- Reduced the length of the Training Manual by 17% by eliminating several pages about the importance of outcomes, names of members of the two NCIL Task Forces, and other materials not directly relevant to the tasks involved during the field test.

## The Second Field Test

Procedurally, the field test in Year Two was very similar to Year One. The Chair of the Task Force recruited 32 Centers, 20 Centers that had participated the previous year and 12 new Centers. They were provided the revised Training Manual and the three revised information-gathering instruments, two 90-minute conference calls for training, two technical assistance conference calls during information-gathering period, and a private listserv for sharing questions and comments.

Centers gathered information from January 10 to early March 2011. This time 28 of the 32 Centers provided information (88%), a noticeable improvement over the 66% rate of Year One. Again, Centers entered information into Survey Monkey, those data were moved into Excel, and the four reports described above were created. Also, Centers again shared their experiences during TA conference calls, by editing the information-gathering instruments, and via a Survey Monkey debriefing survey.

## Lessons from the Second Field Test

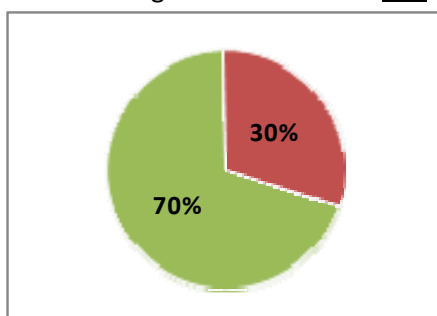
Year Two of the field test produced two different types of findings: First, how Centers are performing on the eight chosen outcomes, and second, about the process of gathering outcome information. Once again, Centers provided helpful insights and suggestions for improving the process even more in the future.

Both types of lessons are presented below, beginning with the findings about outcome performance.

Center performance on eight desired outcomes. Below are Center performances on each of eight desired outcomes. For five outcomes, performance is captured by one indicator. For three outcomes, performance is captured by two different indicators, since one single indicator cannot fully capture the essence of the outcome.

### Desired Outcome #1: Persons with disabilities have skills/ knowledge/resources to support their choices

Seventy percent (70%) of consumers reported they have learned new skills, knowledge, or resources since working with their Center and were able to list one or more of those skills, knowledge, or resources. Most of these 70% of consumers benefited in one of four main ways:



more open-minded")

\* 23% know resources available for them in the community ("learned about heating assistance", "new resources I didn't know existed", "classes and traditional resources")

\* 13% have better social or interpersonal skills ("how to work with people", "confidence as a person with a disability", "to be

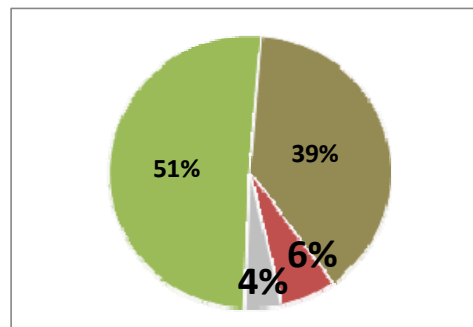
\* 10% are more able to conduct normal day-to-day business ("how to use the computer", "how to get a driver's license", "manage my money")

\* 8% have improved their job-related skills ("interviewing skills", "learning to review time sheets", "how to work in a completely different industry")

Most of the remaining consumers improved their personal care/ADL, obtained new equipment, learned about funding sources, or modified their home.

### Desired Outcome #2: Persons with disabilities are more independent

Fifty-one percent (51%) of consumers reported they have become more independent since working with their Center and were able to describe one or more way they were more independent. Thirty-nine percent (39%) of consumers have the same degree of independence as before, 6% are less independent, and 4% don't know.

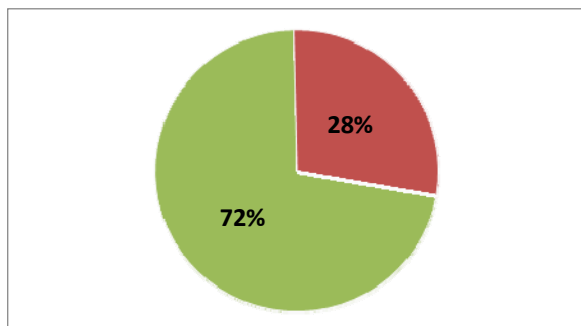


Over two-thirds of these 51% of consumers were more independent in one of three ways:

- \* 31% can do more things for themselves on a basic, practical level ("cooking and cleaning at home more", "I can make important phone calls on my own", "able to bathe on my own now")
- \* 23% have more mobility than before ("taking the bus and keeping a job", "driving", "I took five steps with a walker recently")
- \* 15% have more self-confidence in themselves and a better attitude ("I have more confidence and self-discipline and I feel good about myself", "motivated to get up and out", "willing to try new things")

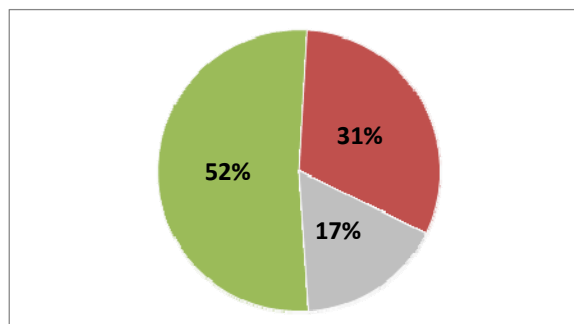
### Desired Outcome #3: Persons with disabilities get the information they need

This outcome concerns persons with disabilities who contact a Center through its I&R activities, and for this outcome there are two separate indicators of interest. First, 72% of I&R callers received the information they needed from the Center. Fifteen percent (15%) of I&R callers did not receive the needed information, and 13% did not remember.



Also, 52% of I&R callers used a new resource they learned about from the Center. Thirty-two percent (32%) of I&R callers did not use a new resource, and 17% did not remember.

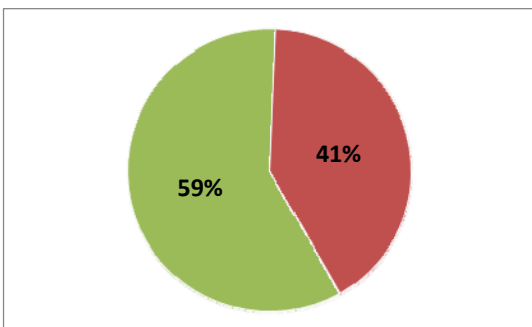
The Task Force believes that the difference between these two findings lends credibility to this outcomes effort. For example, it is encouraging that 72% of I&R callers receive the information they need, but if Centers were skewing data to make themselves look effective, this figure might be closer to 90 or 95%. Also, the difference between the two findings makes



common sense. That is, it seems perfectly understandable that more I&R callers would receive information from the Center (72%) than would use it (52%).

#### Desired Outcome #4: Persons with disabilities advocate for increased community supports

This outcome is also measured by two separate indicators, and each indicator captures a quite different dimension of advocacy. First, 59% of consumers advocated on their own behalf in order to get something they needed (personal advocacy). Most of these consumers advocated for themselves in one of three ways:



\* 19% advocated to get some *services* they needed ("for more training with the white cane", "not in bus riding zone, but I'm riding the bus now", "I met with the YMCA and got a free membership so I can use their installation

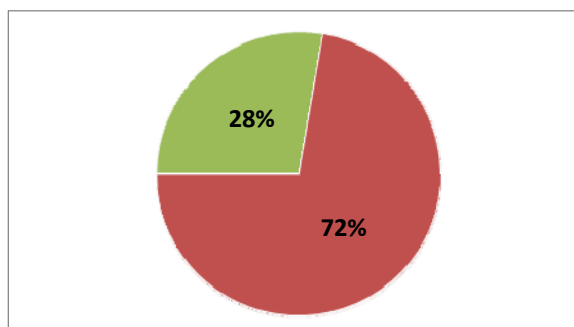
for my physical therapy")

\* 18% advocated to get some *equipment or modifications* they needed ("I asked for a desk that was wheelchair accessible", "I stood up for myself to get a chair", "told landlord I needed a curb cut")

\* 17% dealt directly with an organization to *correct an action or decision* they felt was unfair ("appealed and won wrongful denial of benefits by my insurance company", "advocated to move to a smoke-free apartment, as I had been promised", "school did not want to provide aid")

The remaining consumers advocated for themselves by instructing caregivers how they wanted things done (9%), finding a more appropriate place to live (8%), doing as much independently as possible (7%), stopping something they didn't like (5%), insisting on the personal respect they deserve (5%), lobbying public officials for needed changes (5%), finding an appropriate job (4%), and asking for help when they need it (4%).

At the same time, 28% of consumers advocated to change things in their community that would help other persons with disabilities, too, not just themselves (systems advocacy). Almost half of these consumers advocated for greater physical accessibility ("change church doors to become more accessible with automatic door openers", "placed complaints with City hall about inaccessible buildings and sidewalks", "advocating for people to be treated well while riding public transportation", "asked for lower shelves in grocery store"). The remaining consumers advocated to change or add services, make persons with disabilities more visible, or supported and encouraged other, individual persons with disabilities.

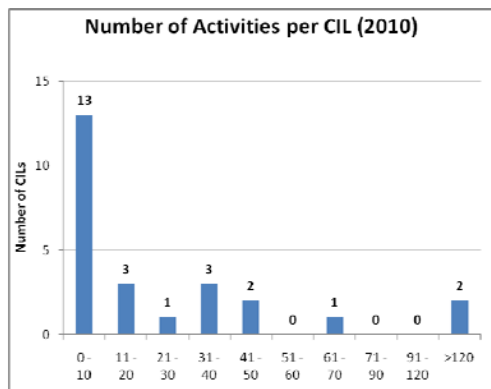


Once again, the Task Force believes that the difference between these two findings lends credibility to the overall set of findings. Neither outcome is enormously positive, so it seems unlikely that the

information is being skewed. And the difference between the levels of personal advocacy vs. systems advocacy (59% vs. 28%), while not ideal, validates common perceptions in the IL field.

#### Desired Outcome #5: Barriers, problems identified

Unlike the first four outcomes, which assess benefits for a Center's consumers or I&R callers, this



outcome relates to the behavior of a Center itself. The particular indicator measures how many activities a Center has conducted "to identify or confirm the primary barriers/problems in the community that prevent persons with disabilities from leading more independent lives." Possible activities might include surveys, public meetings, focus groups, polls, etc.

As the graph shows, Centers vary quite a bit. Thirteen (13) Centers conducted from 0-10 activities, while two Centers conducted more than 120. The overall average is 27 activities, the median is 10, and the most common response (the mode) is only two activities. These findings might

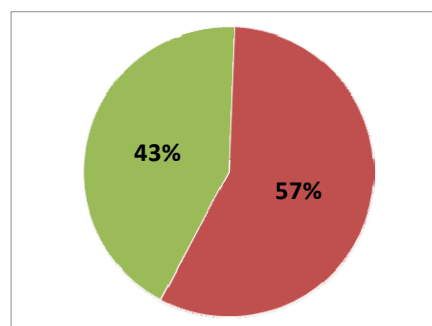
reflect wide variation among Centers, the need to improve measurement procedures, or both.

#### Desired Outcome #6: A consumer agenda for change exists

This outcome also captures the behavior of a Center itself, by asking if a very particular type of workplan exists within the Center – "an explicit and currently-active systems advocacy workplan". The Training Manual provided an example of an acceptable systems advocacy workplan and noted that an acceptable workplan contains:

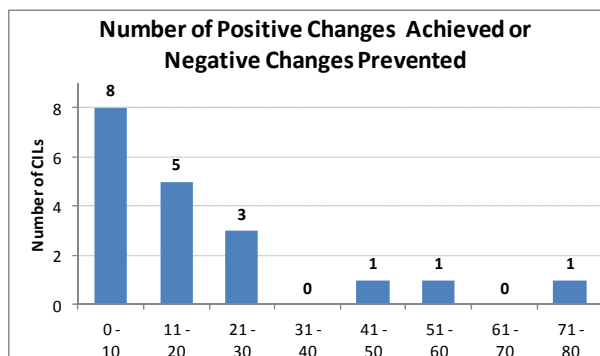
1. *At least one advocacy goal identified by the community*
2. *One or more objectives that address the advocacy goal*
3. *A date when work on each objective is expected to begin, and*
4. *Room to write the date when each objective is met*

Forty-three percent (43%) of Centers participating in the field test have a systems advocacy workplan meeting these four criteria, and 57% do not. Among the "unsuccessful" 57%, however, several Centers have a workplan that meets some of the criteria, though not all four.



### Desired Outcome #7: Decision-makers act on our agenda

Whether they have a systems advocacy workplan or not, Centers are both *achieving positive changes*



and *preventing negative changes* in legislation, policies, practices, or services. To focus Center staff on exactly which types of accomplishments this indicator is measuring, the Training Manual gave examples of each type of change at the local, state, and federal levels

While Centers are less variable on this outcome than they were on outcome #5, they still vary quite a bit. Eight of the 19 Centers providing this information report 0-10 accomplishments, while

one Center reports 71-80 such accomplishments. The average across all Centers is 20 accomplishments, the median is 14, and the most common response (the mode) is seven accomplishments.

Again, these findings might reflect wide variation among Centers, the need to improve measurement procedures, or both.

### Desired Outcome #8: Methods and practices promote independence

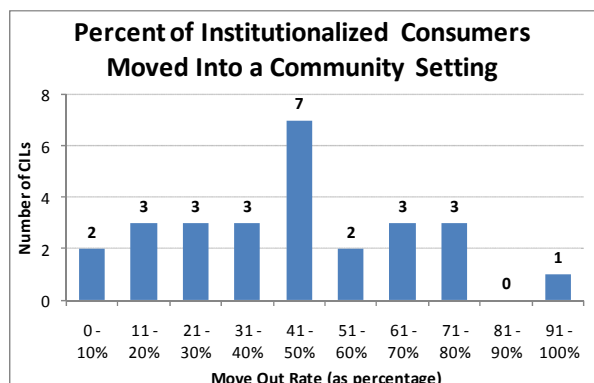
This final outcome was measured by two important, yet also complex, indicators -- each one simple in theory, but difficult in practice. The first indicator captures the percentage of consumers served by the Center who move out of an institution and into a community-based setting -- that is, a Center's "move out" rate. But what constitutes an "institution" and what constitutes a "community-based setting"? The Training Manual defines these terms in this way:

*By "institution", we mean a residence where individual residents do not control and direct one or more aspects of their lives. Institutions may include, but are not limited to group homes, nursing homes, ICF-MR/DDs, boarding homes, some assisted living alternatives, or other small or large congregate living situations.*

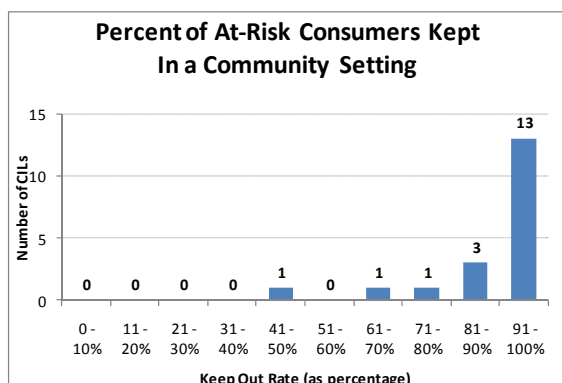
*By "self-directed community-based setting", we mean a residence where individual residents do control and direct every aspect of their lives. Self-directed community-based settings may include, but are not limited to privately owned or rented apartments or homes, some assisted living alternatives, or living with family or friends.*

Overall, Centers moved an average of 30% of their institutionalized consumers into community-based setting during the year. However, Centers vary greatly in the number of institutionalized consumers they serve, and their move out rate also varies. As the figure shows, seven Centers moved 41-50% of their consumers, two Centers moved 0-10%, and one Center moved 91-100%. Other Centers varied along the entire range.

Given this variability among Centers, it is useful to consider the median numbers. A median Center had 30 consumers in institutionalized settings and moved nine of them into the community, for a move out rate of exactly the 30% average.



The second indicator for this outcome captures the reverse of the first indicator – the percentage of consumers served by a Center who remained in a community-based setting despite being at risk of moving into an institution – that is, a Center’s “keep out” rate. Overall, Centers kept an average of 85% of their at-risk consumers out of institutions during the year.



However, once again Centers varied. As the figure shows, 13 Centers kept 91-100% of their at-risk consumers out of institutions, five Centers kept out 61-90%, and one Center kept out 41-50%. Looking at the median numbers, a median Center kept out 62 of

72 at-risk consumers, for a keep out rate of 86%.

Overall observations about these Center outcomes. Looking across all eight outcomes and all 11 indicators, there are three important cross-cutting observations:

1. These findings seem credible. Not only do individual outcome scores seem reasonable, but so also do the relationships between different outcome scores. For example, it seems perfectly reasonable that more consumers would have new skills, knowledge or resources (70%) than would feel more independent (51%), that more I&R callers would receive the information they need (72%) than would actually use that information (52%), and that more consumers would advocate for their own needs (59%) than would advocate for systems change (28%). These patterns of findings give the outcome scores “face validity” – that is, they seem valid “on the face of things”.

Furthermore, the outcomes from Year Two are very similar to those from Year One. While several indicators were changed after the first year, those indicators that were comparable produced quite similar scores for both years. For the eight indicators that were most comparable across both years, the difference in outcome scores ranged from only 2% to a maximum of 9%. This year-to-year consistency suggests that Center outcome measurements are reliable over time.

2. These findings seem to be largely positive. Centers help consumers by providing new skills, knowledge, or resources; making them more independent; helping them advocate, both for themselves and for other persons with disabilities; moving out of institutions; and continuing to live in community-based settings. Centers also help I&R callers by giving them information they need and helping them use that information. Finally, Centers help all persons with disabilities by identifying barriers and problems in their community, creating agendas for change, and working with key persons to achieve positive changes or prevent negative ones. These significant accomplishments are clear reasons for both pride and praise.
3. However, there remains room for improvement. Even while being proud of these accomplishments, the Task Force realizes that each outcome could be improved. For consumers, 30% do not learn new skills, knowledge, or resources; 6% become less independent; 42% do not advocate on their own behalf; 72% do not advocate for an improved system; 70% do not move out of institutions; and 15% of at-risk consumers move into an institution. For I&R callers, 15% do not receive the information they needed and 32% do not use the information they receive. Each of these findings leaves room for improvement, and the Task Force hopes this will spur all Centers, whether participating in the field test or not, to continually improve their effectiveness on these and other important indicators.

The process of gathering outcome information. As noted earlier, Year Two also provided lessons about the process of gathering outcome information. By all accounts, the process ran more smoothly than in Year One. Of the 29 Centers that began gathering information, 28 were able to measure outcomes, Centers raised fewer questions during technical assistance conference calls and on the private listserv, the Chair of the Task Force received fewer private emails asking for help, and the Centers responded more positively to the debriefing survey. Both the Task Force and Centers learned from Year One and were better equipped to measure important outcomes for consumers, I&R callers, and the Centers themselves. This rapid learning bodes well for the future of outcomes in the Center environment.

Detailed findings from the debriefing survey may best explain how participating Centers rate the effort:

- Sixty-two percent (62%) of Centers are confident or very confident that “we were able to get accurate outcome information using these questionnaires and these procedures”. This is 10% better than the 52% figure from Year One, and Centers recommend a continuing effort to improve the wording of the information-gathering instruments.
- Fifty-eight percent (58%) of Centers think it was easy or very easy “to do what we asked [them] to do in this field test”. This is 6% better than the 52% figure from Year One.
- Less happily, 39% of Centers feel the field test required too much time or far too much time. At least this is 5% better than the 44% figure from Year One, and Centers could think of no ways to reduce the time required.
- Seventy-four percent (74%) of Centers think that participating in the field test was valuable or very valuable to their own Center. This is 14% better than the 60% figure from Year One, and the most commonly reported benefit was that Centers gained direct feedback from consumers and I&R callers.



- Sixty-one percent (61%) of Centers thought it was easy or very easy to randomly sample consumers, and 63% thought it was easy or very easy to randomly sample I&R callers. These are 9% and 15% better than the respective figures of 52% and 48% from Year One.
- Sixty-seven percent (67%) of Centers thought it was easy or very easy to ask questions of consumers, and 70% thought it was easy or very easy to ask questions of I&R callers. These are 5% better and exactly equal to the respective figures from Year One.
- Half the Centers (50%) needed to call more than 50 consumers in order to complete 25 successful interviews, and 60% of Centers needed to call more than 50 I&R callers to complete those 25 successful interviews. This workload is not unusual for telephone interviewing, but it was definitely a challenge for Centers.
- Seventy-eight percent (78%) of Centers needed 10 or fewer hours to gather information about their own Centers outcomes, and 93% needed 10 or fewer hours to enter all the outcome information into Survey Monkey. These are 17% better and 3% worse than the respective figures of 61% and 96% from Year One.
- Interestingly, Centers rated all four aspects of Year Two's training and technical assistance to be less "helpful or very helpful" than in Year One. Hopefully this indicates simply that Centers (most of which had previously participated in Year One) were already more prepared for Year Two, and therefore needed these materials less, although it is possible that the Task Force's efforts were simply less successful. Comparative figures for each aspect were the training manual (75% vs. 83%), the original training conference calls before the field test (68% vs. 88%), the ongoing technical assistance conference calls during the field test (36% vs. 71%), and our private listserv for participating Centers (32% vs. 63%).

### **Overall conclusions about outcomes management in the Center program**

Based on two years of field testing, the Task Force draws several overall conclusions about outcomes management in the Center program. First, it definitely *can* be done. When this effort began, no one knew if it would be possible. Centers have gathered two years of outcome information that is quite consistent from year to year, about which they feel confident, and that has high face validity.

Furthermore, Centers themselves *want* to measure their outcomes, despite the effort involved, and they see value to their own Center. The private report comparing each Center's outcomes to the average outcomes of all Centers was very popular, but Centers also felt that the process of contacting consumers and I&R callers was by itself very useful. Centers clearly understand that focusing on outcomes is becoming more essential all the time, not only for reporting performance to various audiences, but also for improving their own effectiveness.

However, there are several unresolved questions as this effort moves forward. In particular:

- Do most Centers *want* to focus on outcomes? Only 32 Centers volunteered for each year of the field test, and only 28 provided a full set of outcome data during Year Two. Do some Centers not

believe in the growing importance of outcomes? Do some Centers believe the task is impossible? What can be done to encourage all Centers to incorporate a focus on outcomes into their philosophy and operations?

- Will Centers want to measure outcomes in their own, unique ways, as opposed to joining in a more *common approach*? If each of the 400+ Centers in the United States develops its own approach to measuring outcomes, the results will be useful for each individual Center, but the lack of consistency will hamper using the findings to advance the wider Center program. The Center program as a whole will benefit only if most Centers can reach a consensus about how to proceed as a group.
- Are these the *most appropriate* outcomes and indicators to measure? From the 16 desired outcomes on the logic model for the Center program, the Task Force focused first on eight outcomes. Do these eight outcomes best capture the essence of the Center program, resonate most to outside audiences, and best help Centers to improve effectiveness? Are these the best 11 indicators to measure progress on these eight outcomes?
- How can the IL field *better conceptualize* some important aspects? In Year One, it was impossible to measure progress on the number of IL goals achieved, and that indicator was deleted for Year Two. Is there a way to better conceptualize this and other issues?
- How can the Task Force *better operationalize* some important concepts? The Task Force looked hard for a workable, operational definition of being “at risk” of being institutionalized. Despite the combined expertise of the Task Force, consultations with other national experts, and extensive research, an acceptable definition was not found. This is a serious gap, not only for outcomes management in the Center program, but for the entire IL field.
- When is the *best time to gather outcome information*? Each field test measured outcomes once, looking back on a nine- or 12-month period. While this was convenient for Centers, some I&R callers, and even some consumers, were interviewed long after their interactions with the Center. The average I&R caller, for example, was interviewed 219 days, or over seven months, after his or her most recent contact with the Center. Perhaps this helps to explain why 13% of these persons don’t remember if they received the information they requested and that 17% don’t remember if they used that information. Perhaps outcome information should be gathered on a rolling basis, perhaps a certain number of days after contact (for I&R callers) or a certain number of days after beginning work together (for consumers).
- How can the *time and effort* required from Centers be reduced? This was the most vexing issue for Centers, especially the number of calls required to complete the 25 successful consumer interviews and the 25 successful I&R caller interviews. As noted earlier, 50% of Centers had to call over 50 consumers, and 60% of Centers had to call over 50 I&R callers. Even if a 3-to-1 ratio of calls to successes may be typical for telephone survey research, this workload places a high demand on Center staff. It seems apparent that we must constantly look for ways to make the process more efficient.

- How can Centers needing *additional training or technical assistance* be identified? Field test Centers were volunteers and each was trained equally and provided with the same support during the process. Even so, some Centers were more effective and efficient than others. It will be important in the future to help needy Centers during the process, but how best to identify which Centers need help?

### **Possible Next Steps**

NCIL, Centers for Independent Living, and the wider IL field will need to decide future directions, but the Task Force sees six logical next steps. We recommend these steps be built into the workplans of all relevant groups.

1. First, and importantly, continue to work closely with RSA to determine what outcomes are appropriate to measure for the Center program, how these outcomes mesh with the current annual performance report, and how any needed changes to Center reporting requirements can be accomplished most effectively and efficiently. At a September 2011 training session on Outcome Measures for Centers, the then-RSA Commissioner affirmed the value of the outcome efforts described in this report and urged all groups to move forward together.
2. Continue to encourage all Centers to focus on outcomes, especially outcomes management and its emphasis on using outcome information within the Center to continually improve the effectiveness of services. NCIL might consider providing Survey Monkey to the Centers as a part of this encouragement.
3. Work with and encourage commercial MIS providers to adapt their data systems to allow local Centers to gather, store, analyze, and report outcome information. Each of the major MIS providers for Centers has expressed its willingness to make these adaptations, and this should occur soon, so as not to slow Centers eager to focus on outcomes.
4. Train other Centers in outcomes management, both in the basic concepts and what the Task Force and participating Centers did and learned in these two field tests. The training session held September 13-15, 2011 in Portland, Oregon was an excellent start in this direction, but more training is needed in order to reach a majority of Centers.
5. Learn from others who are pursuing this same path. The four Centers in Massachusetts, for example, recently conducted their own field test of materials and procedures very similar to these, but with some subtle differences that might prove better or worse in practice. It is important to create a learning community among Centers working with outcomes so that lessons learned, both positive and negative, can be shared.
6. Help Centers learn to use outcome information to improve their services. If Centers fail to use their outcome information, they are simply *measuring* outcomes, not *managing* them. But Center managers are not born with the inherent ability to identify program weaknesses, generate possible improvements, choose the most promising changes, implement those changes, and monitor the effects. These skills are learned, and without them, Centers cannot take advantage of outcome information.