
CBE ID

5120

Title

RTCOC Composite Measure for Home- and Community-Based Services (HCBS) Recipients - Meaningful Community-based Activity

Project

Advanced Illness and Post-Acute Care

Endorsement Status

Not Endorsed

E&M Committee Rationale/Justification

Not endorsed due to no consensus.

Is Under Review

No

Previous Endorsement Cycle

Spring 2025

Steward

Institute on Community Integration

1.0 New or Maintenance

New

1.1 Measure Structure

Composite Measure

1.3 Electronic Clinical Quality Measure (eCQM)

No

1.6 Measure Description

This outcome measure assesses individuals on the activities that are meaningful to them. The target population for this measure is adults with disabilities who receive HCBS or HCBS-like services. This is a self-contained measure that can be administered independently of other RTCOC measures.

The measure is administered through an in-person or video-conferencing interview where an interviewer guides an individual through a series of questions (i.e., items) on the measure. Items

on the measure ask participants about different kinds of activities (e.g., social activities), if that particular activity is meaningful to the participant, if the participant participates in the activity as much as they want to, and if the participant receives enough support to participate in these activities. The 26 items on the instrument are:

- G1. You participate in activities that are meaningful to you.
- G2. You get enough support to participate in activities that are meaningful to you.
- S1. Social activities are meaningful to you.
- S2. You participate in social activities as much as you want to.
- S3. Recently, you have enjoyed doing social activities.
- S4. You get enough help to do social activities.
- S5. Professional activities are meaningful to you.
- S6. You participate in professional activities as much as you want to.
- S7. Recently, you have enjoyed doing professional activities.
- S8. You get enough help to do professional activities.
- S9. Educational activities are meaningful to you.
- S10. You participate in educational activities as much as you want to.
- S11. Recently, you have enjoyed doing educational activities.
- S12. You get enough help to do educational activities.
- S13. Activities that involve physical exercise are meaningful to you.
- S14. You participate in activities that involve physical exercise as much as you want to.
- S15. Recently, you have enjoyed doing activities that involve physical exercise.
- S16. You get enough help to do activities that involve physical exercise.
- S17. Relaxing activities are meaningful to you.
- S18. You participate in relaxing activities as much as you want to.
- S19. Recently, you have enjoyed doing relaxing activities.
- S20. You get enough help to do relaxing activities.
- S21. Everyday life tasks are meaningful to you.

S22. You participate in everyday life tasks as much as you want to.

S23. Recently, you have enjoyed doing everyday life tasks.

S24. You get enough help to do everyday life tasks.

The two items “G1” and “G2” are scored 0 to 3 on a frequency scale with response options:

“Never/Rarely”

“Sometimes”

“Often”

“Almost Always/Always”

The remaining items “S1” through “S24” are scored 0 to 3 on an agreement scale with response options:

“Strongly Disagree”

“Disagree”

“Agree”

“Strongly Agree”

1.7 Measure Type

Patient-reported Experience Performance Measure (PRE-PM)

1.8 Level of Analysis

Population or Geographic Area, Other

1.8a Population or Geographic Area Level of Analysis

Intellectual and/or Developmental Disability, Physical Disability, Psychiatric Disability

1.8b Other Level of Analysis

Home- and Community-Based Services Provider

1.9 Care Setting

Other

1.9b Other Care Setting

Home- and Community-Based Services

1.10 Measure Rationale

Research indicates that in spite of the movement of people with disabilities from institutions to the community, individuals with a wide range of disabilities of all ages participate in community

activities at significantly lower levels than their peers without disabilities (Akyurek & Bumin, 2017; Ginis, et al, 2016; 2021; Rak and Spencer, 2016). A number of barriers to participation have been identified including ableism on the part of the general population, a lack of local resources, poorly accessible transportation, and concern for one's safety in the community (Bezyak, et al, 2017; 2020; Ginis et al., 2016; Levasseur, et al, 2015; Vasudevan, et al., 2016). Other challenges people with disabilities face include the costs of engaging in many community-based activities and their personal financial situations (Hästbacka, et al., 2016). This state of affairs was exacerbated by the COVID-19 pandemic (Courtenay & Perera, 2020) from which access to meaningful activities, especially those that are community-based activities has yet to recover.

In addition to the need to measure outcomes in this area in order to establish compliance with the HCBS Settings' Final Rule (2014), an understanding of the extent to which HCBS recipients are able to engage in activities that they find meaningful is critical for providers given the association between activity engagement and other important outcomes. These include both physical and mental health (Ginis, et al, 2021; Rowley, et al., 2018; Theis et al 2021; Zar, et al., 2018).

Fortunately, many of the challenges to meaningful activity faced by people with disabilities are actionable and a number of facilitators have been identified that research clearly indicates have the potential to enhance access and engagement in meaningful activities (Aherne & Coughlan, 2016; Alesi & Pepi, 2015; Di Domenico, et al 2022, Dixon-Ibarra et al., 2016, Mahy et al., 2010, Temple & Stanish, 2011). Providers with access to information with respect to the meaningful activity outcomes and needs of the people they serve therefore have the opportunity to engage in quality improvement in this area.

Further rationale for why measured entities should report outcomes related to the **Meaningful Activity ID** are grounded in what people with disabilities, their families and other stakeholders indicate is the relevance of the measure, and the potential of this measure to reveal important aspects of quality of life experienced by individuals across disability groups.

Of critical relevance related to the inclusion of this measure is its *Importance and Meaningfulness* to people with disabilities and other stakeholders. Section 2.6 of this submission for further information as to the multiple ways that this was determined with respect to the meaningful activity construct that formed the basis of the measure being submitted.

In addition to its relevance to people with disabilities, the construct of meaningful activity in both the community and elsewhere is also referenced in policy and legislation underlying home and community based services. In 2014, the Home and Community Based Services (HCBS) Settings' Final Rule came into effect. This CMS policy stipulates that in order to receive Medicaid reimbursement for providing HCBS, it must be demonstrated that these services are delivered to individuals with disabilities in such a way that opportunities for people to have access to the benefits of community living, including receiving services in the most inclusive community-based settings, are maximized. This emphasis on receiving services in the most inclusive settings is aligned with states' efforts to meet their obligations under the ADA and the Supreme Court Olmstead decision (527 U.S. 581, 1999).

Measurement at the agency and MCO level requires assessment tools that go beyond those used for compliance at the state and national level. They need to assess outcomes based on the NQF framework as this framework has been validated by multiple stakeholder groups, including people

with disabilities, family members, policy makers, researchers as well as service providers. Measures need to be sensitive to the needs of provider organizations and MCOs involved in the delivery of HCBS as well as the populations they serve (i.e., performance scores directly measure outcomes toward which an organization is attempting to achieve service quality improvement or demonstrate improvement in recipient outcomes). Measures should possess the capacity to longitudinally track progress on key indicators/outcomes that are within the possibility of providers to improve.

Although a number of instruments are available that have attempted to measure meaningful activity outcomes of persons with disabilities, only a limited number have included items that attempt to solicit information about how effectively the system supports the individual's involvement in such activities. In addition, few currently available instruments have demonstrated adequate reliability and validity and can be used across different disability populations. In order to be most useful in quality improvement and efforts, a measure performance score for *meaningful activity* must include information related to (a) the importance/value to respondents of each type of activity, (b) the levels of enjoyment/satisfaction individuals derive when they participate in each type of activity, (c) the degree to which HCBS recipients are able to engage in activities that are meaningful to them both at home and in the community to a degree that meets their unique needs and preferences, (d) the extent to which these activities are inclusive or of a segregated variety, (e) whether people are able to engage in these activities with preferred others, (f) the degree to which people receive sufficient support to engage in the types of activities that meet their needs and (g) the extent to which the support provided through HCBS encourages participants to be as independent as they are able?

The *Meaningful Activity* measure developed as part of the RTC/OM was designed to target outcomes of people with disabilities at the agency level and be sufficiently sensitive to changes in policy and services that could document improvement in quality of both services and HCBS recipient outcomes. This level of measurement needs to be more granular than the level focused on compliance and requires measures able to demonstrate reliability, validity, accuracy, and sensitivity at the agency level with specific groups of HCBS recipients each organization serves. The measure is also constructed to be person-centered, taking into consideration that personal preferences/interests and accordingly the importance people assign to taking part in different activities vary considerably and must be taken into consideration in performance scores.

Information/data available based on performance scores on this IDM have the potential to provide support agencies with a variety of information that can be used to (a) document overall service quality and facilitate policy and/or programmatic changes needed as part of quality improvement efforts, (b) identify specific aspects of the meaningful activity subdomain where performance is less than desirable as well as those areas in which the agency is supporting exceptional outcomes, (c) longitudinally track changes that occur in service quality and outcomes, and provide families and persons with disabilities with information they can use to help make informed decisions as to which providers they desire to provide services to their family members with a disability.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

1.13 Data Dictionary

Attached

1.13a Attach Data Dictionary

[RTCOM Codebook-and-Scoring-Scheme - Meaningful Activity + ServicesSupports.pdf](#)

1.14 Numerator

The measure focus for the performance measure outcome is the number of individuals who have a composite score within the measure's possible inter-quartile range (IQR). An individual's composite score is derived by summing the responses to 26 items on the Meaningful Community-based Activity instrument. (See the response to section 1.6 for details on item scoring.) An individual's composite score can range from 0 to 78 where higher scores indicate a participant's greater overall involvement in activities that are meaningful to them. The IQR is 20 to 58 and the numerator is the number of individuals whose composite scores fall in this range.

1.14a Numerator Details

The numerator is calculated using a tabulation of composite scores from individual respondents. These individuals belong to the relevant population as defined in the denominator for this measure. Other details such as time period for data collection are equivalent to the denominator definitions and will be discussed in section 1.15a.

A composite score for an individual respondent is obtained by administering the Social Connectedness instrument to a respondent then calculating a sum score from the responses to all items on the instrument. This results in a composite score for one individual. See the attached data dictionary (1.13a) for a list of items, response options, and response scoring. An individual composite score (i.e., sum score) for the Social Connectedness measure can range between 0 to 40.

Included in the numerator are all individual respondents included in the denominator who have composite scores within the measure-derived inter-quartile range (IQR) for the measure. In other words, composite scores that fall within the middle 50% of possible scores that can be obtained for the measure. For the Social Connectedness measure this range is 10 to 30.

It is recommended that these calculations be performed prior to scaling of scores to be on a different, public-facing metric, e.g., T-scores.

1.15 Denominator

The target population for this measure is individuals receiving HCBS who are at least 18 years old and have a primary diagnosis of either: intellectual and/or developmental disability, physical disability, and psychiatric disability. Respondents must be able to complete an interview either independently or with assistance (e.g., support staff).

1.15a Denominator Details

The target population for this measure is individuals receiving HCBS who are at least 18 years old and have a primary diagnosis of either: intellectual and/or developmental disability, physical disability, and psychiatric disability. Respondents must be able to complete an interview either

independently or with assistance (e.g., support staff). The available research supports measure administration every four to six months.

1.15b Denominator Exclusions

None

1.15c Denominator Exclusions Details

None

1.15d Age Group

Adults (18-64 years), Older Adults (65 years and older)

1.16 Type of Score

Composite scale

1.17 Measure Score Interpretation

Better performance = Higher score

1.18 Calculation of Measure Score

The measure score for individual recipients is calculated by the sum of item responses on the instrument. See section 1.6 and the attached Data Codebook in 1.13 for the items and scoring codes on this instrument. Stratification does not modify the calculation of measure scores. Note that the measure score interpretation in section 1.17 is only applicable to this individual measure score.

Scores at the accountable-entity level are calculated using individual measure scores and incorporating the numerator and denominator criteria provided in sections 1.14/1.14a and 1.15/1.15a. For a calendar year, accountable entities will obtain individual measure scores for eligible HCBS recipients that they serve. The numerator is the number of individuals who obtained a score in the measure-derived interquartile range (IQR; see 1.14/1.14a). The denominator is the total number of eligible HCBS recipients that were assessed. The accountable entity score is this numerator/denominator ratio, which will be a proportion.

Each individual HCBS recipient should only be counted once in the numerator and denominator. In other words, an HCBS recipient with more than one assessment within a calendar year will still be only counted once for reporting purposes. For these individuals, use their most recent individual measure score that still falls within the calendar year.

A higher score indicates that there is a greater performance gap on measure outcomes for recipients at a given accountable entity. A higher proportion also suggests that the provider serves a wider population with greater variability in service needs. We have purposefully avoided an accountable entity score interpretation that is based on high (or the highest) individual measure scores as this would provide perverse incentives for those using our instruments. We also

do not recommend higher or lower accountable entity scores be interpreted as “better” or “worse”. Rather, accountable entity scores should be accurate and informative.

1.19 Measure Stratification Details

See data dictionary attachment.

The data dictionary contains item and scoring schemes for questions that ask respondents about their Services Needs. These questions are found on the RTC/OM’s Demographic survey. This survey is appended to the instrument information in section 1.13a Attach Data Dictionary.

The Service Needs questions are used to develop a classification variable of a respondent’s functional disability (see section 5.4.2). The functional disability variable is defined as the overall level of services and supports a respondent is currently receiving. During field testing of the RTC/OM instruments, this collection of items was found to be superior in identifying level of disability compared to items that directly asked participants about their functional difficulties (see question 8 in the Demographic survey, attached in section 1.13a). In essence, most participants indicated low levels in functional difficulties despite varying levels of service needs. This indicates that services are effective: they reduce or eliminate the difficulties a person experiences in various areas of their life. As such, level of current service needs is a better indicator of functional disability.

Scores on the service needs items are summed to create a composite score of service needs for a respondent. Higher composite scores indicate greater service needs. Cut points are used to group individuals into functional disability categories:

- 0 - 10 = Few or no services & supports
- 11 - 20 = Moderate services & supports
- 21+ = Intense services & supports

Our research team is cognizant that polychotomizing a numerical scale into a few discrete categories is not best practice, as it reduces the amount of information (i.e., variance explained) in the outcome of interest. However, we chose this categorical approach to make allowances for those who have less technical expertise and want an easy-to-understand display of the relationship between level of functional disability and service needs.

1.20 Types of Data Sources

Patient-Reported Data and/or Survey Data

1.20c Format: Patient-Reported Data and/or Survey Data

Non-digital

1.21a Data Collection Tool URL(s)

<http://example.com>

1.21b Attach Data Collection Tool(s)

[RTCOM Instruments - Meaningful Activity + Demographics.pdf](#)

1.22 Proxy Responses

No

1.23 Survey Respondent

Other

1.23a Other Survey Respondent

HCBS Recipient

1.24 Data Collection and Response Rate

The Meaningful Activity Measure is reported by HCBS recipients and administered either in person or through a HIPAA-compliant remote video conferencing platform. Currently, the instrument is available only in English; however, an interpreter can be used during the interview to improve accessibility. Response rates may be improved by offering accommodations that increase accessibility, such as environmental modifications for individuals with disabilities and visual displays of response scales.

Given that accountable-entities (i.e., HCBS providers) can vary significantly in the number of individuals that they serve and services that they provide, we do not require a minimum response rate. For example, some providers may only serve fewer than 5 individuals. However, for organizations that serve recipients with a wide range of service needs, then the sample used to calculate response rates should be representative of these service needs.

1.25 Data Source Details

N/A

1.26 Minimum Sample Size

There is no minimum sample size that is needed to calculate a performance score at the individual HCBS recipient level. Performance scores for all of the IDMs submitted are based on composite scale scores and it is these scores that are reported.

Given that the composite measures developed were intended to be used at the provider level to document service quality and HCBS recipient outcomes, it is unlikely that the measures would be used to represent the quality of services received and outcome experienced by a single individual. As a result, while performance scores could be calculated on the basis of a single individual they would have little meaning in that isolated context. Focus group results and other discussions we have had with HCBS providers indicate that the most likely future use of the measures will be to (a) document the extent to which providers and sites within provider networks meet established benchmarks, (b) determine the extent to which quality improvement efforts have had a significant

impact on service quality and HCBS beneficiary outcomes in targeted areas, and (c) potentially be used in the establishment of “report cards” for agencies within and between provider networks that can be used by external funders and families to determine programmatic strengths and areas in need of improvement and to facilitate informed decision-making when people with disabilities and/or their family members are looking to contract for services.

When used in this manner, sample sizes deemed adequate for analysis and interpretation of the submitted measures will depend on a variety of factors including, but not limited to the questions one desires to answer, comfort levels in committing Type I (false positive) and Type II (false negative) errors, the sampling methods used (i.e., probability vs. non-probability sampling), the target population and its attributes. Attributes of HCBS samples include a wide variety of factors that impact sample size and include client demographic variables such as disability type, intensity of support needs and their variation within the population of interest, geographic location, the size of the population, the effect size one expects and a variety of other features of clients and the environment that could be relevant for answering a provider’s evaluation questions.

The University of Minnesota’s Institute on Community Integration, home of the RTC/OM has long-term history of working with HCBS service providers on a wide variety of evaluation projects that have included the use of measures similar to the ones we are submitting for review. These experiences suggest that if a provider desires to use performance scores aggregated across individuals to determine *changes* within that provider organization, a representative sample (adjusted for population size) that includes an absolute minimum of 35-50% of HCBS beneficiaries receiving services is needed to draw valid conclusions regarding quality improvement efforts.

2.1 Attach Logic Model

[EM-Logic-Model-Meaningful Activity \(3\).docx](#)

2.2 Evidence of Measure Importance

General Importance

Importance is defined as the relevance of measures to the lives of people with disabilities who receive HCBS and the potential of these measures to reveal important aspects of quality of life experienced by individuals across disability groups.

Importance was determined via three sources/processes. First, we began by considering all the domains and subdomains identified by the National Quality Forum’s (NQF) Framework for Home & Community Based Services Outcome Measurement (2016). The framework for HCBS outcome measurement was developed by a panel of NQF experts with the goal of identifying key areas to measure in order to be able to track the effectiveness of HCBS services. The final recommendation of the panel included 11 domains with 2-7 subdomains within each domain.

Next, we engaged stakeholders on whose lives HCBS has an impact in a series of feedback and

planning groups. Stakeholders included: (1) Persons with disabilities (100 participants), including individuals with intellectual disability and developmental disabilities, physical disabilities, traumatic brain injury, mental illness, and age-related disabilities; (2) family members (84 participants); (3) providers (89 participants), and (4) program administrators/policy makers (47 participants) for a total of 320 participants in 58 small groups conducted nationally. Stakeholders took part in a participatory planning and decision-making (PPDM) process in which participants weighted each domain and subdomain of the NQF framework on a scale from 0-100 based on their perceived importance in determining the HCBS outcomes and quality.

As a third step in the process, we solicited the input of measurement experts in disability-related fields. Two groups of experts rated all subdomains, including the new subdomains of employment and transportation, in terms of feasibility, usability, and importance. These expert ratings were used in conjunction with stakeholder weighting from PPDM groups to narrow our development process to nine NQF subdomains on which to initially focus our work.

Specific Importance

Stakeholders who took part in the PPDM groups indicated that meaningful activity was important to measure as part of the larger domain of community inclusion. In addition, meaningful activities have important quality of life implications due to the importance of including people who use HCBS into their communities in accordance with their personal preferences. The meaningful activity measure concept aims to address the need for a sound measure to assess the level of engagement of people who use HCBS in desired activities.

To define and understand the construct at stake we primarily used the NQF definition and approach. Meaningful activity is defined by the NQF as the level to which individuals who use HCBS engage in desired activities (e.g., employment, education, volunteering, etc.). Because the employment measure concept is under separate development, we excluded that concept when developing the meaningful activity measure concept.

Meaningful translates to the congruity with one's value system and needs, its ability to provide evidence of competence and mastery, and its value in one's social and cultural group (from Engagement in Meaningful Activity Survey, Goldberg et al., 2002). While the word meaningful is correct and valid, the term desired, favored by NQF, seems to be more appropriate considering the person centeredness of this measure concept. Activity, according to the World Health Organization, is "the execution of a task or action by an individual." Using the term(s) participating/participation may be a suitable addition for our purpose of targeting specific groups of activities within a more conversational format.

Research has indicated that participation in meaningful activities by people with different types of disabilities is critical to their health and wellbeing (Sellon, 2021) and quality of life (Oh et al., 2021).

(A complete reference list is provided as a supplemental attachment in section 7.1.)

2.3 Anticipated Impact

The intent behind the Meaningful Activity measure is to impact the outcomes of meaningful activity in the context HCBS excluding employment. Meaningful activities are those activities that are important, enjoyable, and/or valuable to the individual and the activities can be instrumental or functional, social and/or cultural, and recreational or leisure-type with different levels of physical demand.

The outcomes we intended to impact with the meaningful activity measures include the HCBS service recipients' engagement in activities at home and in the community to the extent they desire (Qian et al., 2015); expansion of natural supports and community connections as a result of participation in meaningful activities (Sanderson et al., 2024); and services to support people with disabilities to engage in meaningful activities are delivered with greater efficacy, timeliness, and quality (van Herwaarden et al., 2025).

No unintended consequences have been identified.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

2.4 Performance Gap

Measures are being submitted for initial endorsement.

2.5 Health Care Quality Landscape

Most of us, regardless of whether we have a disability, desire to engage in preferred activities with the people we desire, at times that are most convenient to us, and to the degree that we desire. The extent to which people with disabilities are able to live the types of lives they desire is often far more dependent on the availability and effectiveness of the paid and unpaid support they receive from others than for the general population. The capacity to monitor the extent to which people with disabilities are able to engage in activities that reflect their preferences to a degree comparable to that of individuals without disabilities is critical if we are to understand the extent to which community-based services are doing what they are intended to do. At one level, performance measures are needed to ensure compliance with federal and state regulations governing HCBS as it relates to community inclusion. Performance measures, however, are also needed that are person-centered, longitudinal to assess outcomes associated with various aspects of life as well as the quality of support service recipients receive. These measures need to be sufficiently sensitive to change that the impact of policy, funding, and programmatic changes on the outcomes people experience can be determined over time. They would also preferably have the capacity to be used with different disability populations who receive community supports.

As a result of the variety of HCBS waiver programs and diversity of users, measurement of the quality of supports that the recipients receive and the meaningful activity outcomes they experience is far from straightforward. A nuanced approach needs to be taken that is responsive to a wide variety of personal and contextual factors. This process needs to be decidedly different than that currently used in healthcare contexts due to the dissimilarities in the constructs

measured. Unlike many outcome measures related to health (e.g., the number of urinary tract infections or falls experienced by a person, blood pressure, etc.) outcomes associated with meaningful activities are both more complex and difficult to assess. A second set of critical contextual factors for which one needs to account are the policies and regulations under which HCBS is implemented which vary significantly between states in the U.S. At the same time, in order to be confident that performance measures associated with HCBS adequately assess both the quality of services and the outcomes people with disabilities experience, data are needed with respect to their reliability, validity, and sensitivity to change. Indicators of quality and unmet support needs as directly perceived by service recipients must be considered paramount when developing, administering, and interpreting results based on these measures.

Over the past twenty years there has been great interest in assessing the degree to which people with disabilities are able to engage in activities that are personally meaningful. During this period, CMS has championed the development and maintenance of the HCBS Quality Measure Set (QMS). The QMS is intended to promote more common and consistent use within and across states of nationally standardized quality measures in HCBS programs, create opportunities for CMS and states to have comparative quality data on HCBS programs, drive improvement in quality of care and outcomes for people receiving HCBS, and support states' efforts to reduce disparities in their HCBS programs. The QMS is intended as a resource for states and thus focuses on the compliance level. The approach to measurement taken, however, is decidedly medical in its orientation. As such, "performance measures" are most often conceptualized as single items. While this might be appropriate to measure discrete healthcare outcomes, most psychometricians would argue that it is inadvisable to use when one is attempting to measure *latent* variables including the perceptions of individuals as to whether they have engaged in activities that are meaningful. In all but a few cases, latent variables should not be measured with single items because single-item measures all too often lead to inaccurate representation of the latent construct and limit the ability to assess measurement error. Using composite measures composed of multiple items provides more robust evidence of construct structure, allows for estimation of measurement error, and enhances the overall validity and reliability of the measure.

The National Core Indicators (NCI/NCI-AD; HSRI <https://www.nationalcoreindicators.org>) is currently the most widely used tool in the U.S. for the assessment of outcomes associated with the receipt of HCBS. The instrument was developed and validated as a state-level compliance measure and does an excellent job when used at that level. It is not, however, intended to be used at the provider or individual level for quality improvement, service plan development, and/or outcome assessment. In addition, although the NCI includes indicators in a variety of areas, it is intended to be administered (and was validated) at the instrument level as opposed to on an indicator-by-indicator basis. Users are therefore required to administer items related to all indicators as opposed to only those in which there is a specific interest. It should also be noted that in research undertaken at the University of Minnesota's Institute on Community Integration by Ticha and colleagues, the NCI indicator "meaningful activity" did not perform well.

CQL's Personal Outcome Measures (CQL, 2017) is one of the better developed and validated HCBS Outcome tools and part of a commercially available system of assessment and quality improvement. It has been validated with a much wider variety of people with disabilities than the NCI and possesses good psychometric properties. However, the instrument is time consuming with respect to administration (715 items; 12) limiting its feasibility for many providers. In

addition, the CQL-POM, as part of a quality improvement package, is proprietary and expensive to use.

A third approach to outcome assessment in the human services field that has been championed by CMS is the HCBS CAHPS Survey. The CAHPS is a questionnaire with sixty-nine core items developed for measuring the experiences of people with disabilities who are HCBS recipients. The CAHPS, unfortunately, currently has limited data available with respect to its validity or reliability. Internal consistency reliabilities for seventeen of its nineteen measures including those related to meaning activity fail to meet even the most basic criteria for psychometric acceptability, there are serious questions about the representativeness of the sample used for the field study as well as the evidence presented to support validity, and in a number of indicator areas, there appears to be a ceiling effect with the overwhelming majority of respondents indicating the highest possible level of service quality or personal outcomes (Nyce, et al., 2020).

In addition to the individual shortcomings of the most widely used HCBS outcome measures, there are additional limitations that cut across these instruments as well as other HCBS outcome assessment tools that contribute to the need for development of new measures. The first of these entails the small percentage of items included in HCBS outcome measurement instruments that meet the criteria for person-centeredness. Recent decades have seen a growing focus on providing HCBS in a person-centered manner thereby supporting outcomes that are both important for and to the person. No longer is it sufficient to focus services on what is important *for* the person. Rather, supports must reflect both what is important for and what is important to the person (Smull, 2017).

In addition to the CMS/HCBS system's move toward person centered service provision, there are legal and compliance motivations within the HCBS environment that support the need for measurement that is person-centered. In 1999 the U.S. Supreme Court ruled in *Olmstead v. L.C.* that unjustified segregation of persons with disabilities constituted discrimination and was in direct violation of title II of the Americans with Disabilities Act. Under the *Olmstead* decision (1999), as well as the HCBS Final Settings Rule (2014) states in the U.S. are now obligated to provide services for people with disabilities in the most inclusive community settings possible as well as support them to achieve desired life outcomes. To fully measure the effectiveness of programs that provide services and supports in meeting *Olmstead* and recent CMS requirements related to HCBS, *a person-centered approach to measurement is needed*. The approach needs to emphasize the degree to which the outcomes experienced by HCBS recipients match their needs and preferences and move them forward in achieving desired life outcomes. A simple counting of the number of time a person has gone shopping or out to eat in the community is simply not a person centered measure.

HCBS outcome measurement has not kept pace with advancements in person-centered thinking as it relates to providing supports to people with disabilities. The concept of person-centered outcome measurement has been inadequately defined and is frequently misunderstood, including by those in the measurement field. A study of 140 outcome measures used with HCBS populations (RTC/OM, 2017) found that only 36% of the items included in these tools were person-centered in nature. Although some outcome measures (e.g., the CQL-POM) are more person-centered than others, the overall results of this study clearly indicate the need for approaches to assessment that place greater priority of assessing outcomes within the context of what is most important to individual.

A second shortcoming that cuts across the majority of HCBS outcomes measures is the lack of evidence that they are sufficiently sensitive to change over time that they can be used in a longitudinal manner. Some developers, such as HSRI (NCI-ASC/NCI-AD) explicitly state that their measures including *meaningful activity* are not intended to be used longitudinally. Others (e.g., CQL, CAHPS) have yet to provide sufficient evidence that, when used in a longitudinal manner, their measures are sufficiently sensitive to change that they can be used as evidence of the effectiveness/efficacy of quality improvement efforts, changes that take place in a HCBS recipient's life, disability policy or funding or as part of value-based payment systems.

A third reason to think about the development of new outcome measures for HCBS emanates from the resources needed to administer measures at a time when the human services field is experiencing serious workforce shortages. All of the tools mentioned above are intended to be administered in their entirety as full instruments. They are neither modular in format allowing for administration focused on only one or a few indicators, nor tiered and able to provide both a quick general overview of indicators as well as a more in-depth assessment needed for having utility at the provider level.

A final overarching rationale for considering the need and development of new measures is that the best developed and most well-researched measures that are currently available in the field are proprietary and part of measurement systems. States as well as large providers typically have the funding to pay for the use of these tools. Provider agencies, especially those of the small variety, however, often do not. As a result there is a need for measures at the provider level that are (a) able to be used a little to no cost, (b) person-centered, (c) of a composite nature with the ability to assess latent constructs, (d) based on recent theory and research pertaining to the outcome domains and subdomains assessed, (e) easily scored and interpreted, and (f) sufficiently sensitive to change over time so that they can be used longitudinally.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

2.6 Meaningfulness to Target Population

Stakeholders, particularly those who have disabilities and their supporters, should be at the heart of measurement development. In order to ensure that home and community-based services (HCBS) outcome measures are of high quality, the measure development process must include input from stakeholders and most importantly the intended population with which the measures will be used. Furthermore, we contend that the unique challenges associated with measuring outcomes among members of diverse populations of HCBS recipients requires strong stakeholder involvement throughout all stages of development. This process has been affirmed by NIDILRR, ACL, and the Centers for Medicare and Medicaid Services (CMS, 2019). Using a sound HCBS outcome measurement framework that has evidence of content validity provided by stakeholders including people with disabilities; putting all measures developed through multiple expert panel reviews; undertaking cognitive testing with people with a variety of disabilities are necessary strategies or processes in which one must engage to ensure for quality measure development.

The target populations for the submitted measures includes people with intellectual and developmental disabilities, psychiatric disabilities, physical disabilities, and TBI/ABI as well as age-related disabilities. Although the large majority of people with such disabilities possess the

capacity to articulate their thoughts and feelings about the outcomes in question and their importance, it must be recognized that some do not. The intensity of their support needs may be such that they experience difficulty understanding questions and articulating their thoughts and feelings. We often refer to these individuals as the “forgotten ones.” Not only are they at elevated risk of experiencing poor services and outcomes but most HCBS performance measure programs are not set up to reflect their experiences. The RTC/OM development team therefore instituted a process in which not only HCBS recipients (as opposed to patients) provided input into the measure development process, but other stakeholders were included as well. These consisted of family members of people with the above noted disabilities (who are often direct caregivers on either a part or full-time basis), paid caregivers, and HCBS program administrators who in the end will be responsible for using performance measure data to improve both services and outcomes. At multiple steps along the measure development process, these individuals were consulted and their input incorporated into the measure development process. All of the measures submitted for review were based on:

1. The results of a national content validation study of the National Quality Forum’s HCBS Outcome Measurement Framework using a participatory planning and decision making process.
2. Utilization of multiple Technical Assistance Panels (TEPS) that included people with disabilities as well as members of other stakeholder groups
3. Input from an RTC/OM Center Advisory Committee composed of people with a variety of disabilities in addition to other stakeholder groups, and
4. Extensive cognitive testing of measure items and response options, and
5. Vetting of measures by representatives of provider agencies

PPDM Process. Of critical relevance related to ensuring that HCBS recipients for whom the measures under development were intended viewed them as important, we began by considering all the domains and subdomains identified by the NQF Framework for Home & Community Based Services Outcome Measurement (2016). The initial framework for HCBS outcome measurement developed by the NQF covered 11 domains as well as 2-7 subdomains within each domain. As our initial step in determining importance and relevance, University of Minnesota RTC/OM staff engaged stakeholders from 29 states on whose lives HCBS has an impact in a series of participatory planning and decision-making (PPDM) groups. Stakeholders included: (a) people representing with disabilities (100 participants), including individuals with IDD, physical disabilities, traumatic brain injury, psychiatric disabilities, and age-related disabilities; (b) family members of people with the above noted disabilities (84 participants); (c) representatives of HCBS provider agencies (89 participants), and (d) state level HCBS program administrators/policy makers (47 participants) for a total of 320 participants who took part in 58 small (4-6 person) PPDM groups.

The PPDM process initially included meeting with homogeneous (with respect to disability and stakeholder type) stakeholder groups and providing them with an opportunity to evaluate the original NQF framework. They were provided with the opportunity to add to it, remove domains and/or subdomains they believed were not important, and then stipulate which personal outcomes and service characteristics were most important to measure. Following stakeholders reaching consensus with respect to the domains and subdomains of the original framework they wanted to

add or remove, members of each group took part in a process in which they first independently assigned importance weights for each domain and subdomain of the original NQF framework on a scale from 0-100 (or 0-10 for persons with cognitive challenges) based on their perceived importance in determining the HCBS outcomes and service quality. As part of the PPDM process, stakeholders then discussed their weightings first at the subdomain level and later at the domain level examining why people in their group assigned the importance weightings they did. Stakeholders were then given the opportunity to assign a second set of importance weightings taking into consideration what they had heard during their discussion. Overall, results from PPDM groups indicated a high degree of stakeholder support for inclusion of a meaningful activity measure with a mean importance weighting across groups of 92.78/100 (SE = .73).

Technical Expert Panels. As a second step in the process of developing measures, we solicited the input of people with lived experience with disability (N = 9) as well as measurement and content experts in disability-related fields (N = 12) for a series of technical expert panels (TEPs). Four TEP groups were formed and initially asked to rate the importance of all subdomains in the NQF framework that received the highest importance weightings across stakeholder groups on a 1-5 point Likert-type scale with respect to their feasibility, usability/utility, and importance, as well as provide an overall score for the subdomain. Across stakeholder groups the meaningful activity subdomain received the following weighting with respect to its *feasibility* (Mean = 4.5/5.0), *usability/utility* (Mean = 4.0/5.0), *importance* (Mean = 4.8/5.0), as well as *overall score* (Mean = 4.5/5.0). This information, in combination with PPDM results, a systematic review of the literature, and analysis of existing HCBS outcome measures indicates both the meaningfulness of the meaningful activity construct and perceptions among stakeholders that it is important to measure.

At follow-up TEP meetings following the initial development of *items* for those subdomains that were selected to be a focus of RTC/OM work, TEPs engaged in a similar process using 1-4 point Likert-type rating scales to rate the *relevance*, *importance*, *accessibility/understandability*, and *accuracy* of *each item* developed for the meaningful activity measure under construction. Based on this information low-rated items were jettisoned and replaced with new ones. Mean scores for *meaningful activity* items were as follows: *Relevance* (Mean = 3.44/4.00), *Importance* (Mean = 3.47); *Accessibility/Understandability* (Mean = 3.21); and *Accuracy* (Mean = 3.2).

RTC/OM National Advisory Group. In conjunction with the results of PPDM groups and TEPs, ongoing input was also solicited from a national advisory group of HCBS stakeholders composed of 14 individuals including 5 with lived experience with disability. The RTC/OM Center Advisory Committee provided valuable feedback not only with respect to the item content of the measures under development but measure administrator training content and the medium (live versus virtual using HIPAA compliant version of Zoom Meeting) through which interviews would be conducted and the approach to cognitive testing that would be used after item refinement.

Cognitive Testing. Cognitive testing (CT) is designed to obtain direct input from respondents to verify their interpretation of items and the words of which they are composed to ensure that these match the developer's intent (Ericsson & Simon, 1980; Willis, 2005; Willis, et al., 1991). It is an essential step for ensuring the meaningfulness of items and scales, improving item accessibility (Kramer & Schwartz, 2017) as well as contributing to the validity of measures (Castillo-Díaz & Padilla 2013). RTC/OM staff used a cognitive testing strategy referred to as the "Think Aloud Method" to address the core cognitive components of item responding as included in the

Cognitive Aspects of Survey Methodology (CASM) model: comprehending the item, retrieving the information needed to answer the item, making a judgment, and reporting a response (Tourangeau, 1984; 2018). This approach provided yet another way to involve people with disabilities in the measure development process and allowed developers to modify items to maximize understanding and meaningfulness.

Focus Groups with Potential Measure Users. It is essential that people with disabilities have multiple opportunities to provide input into the measure development process and resulting measures. The performance measures under development, however, also needed to undergo vetting by representatives of provider agencies. If this group did not view the measures as being technically sound, having utility, and being feasible to administer and use, the work of RTC/OM would be little used. After final performance measure refinement based on the results of piloting and field-testing the final measures under submission were presented to two separate groups of potential users utilizing a focus group format. Participants were recruited from a large human service organization in Minnesota that provides residential, home health, and employment services, and a statewide network of human services providers located in Michigan (Total N = 23). Measures were initially shared with participants several days prior to scheduled focus groups. Groups were initiated with RTC/OM providing background information and answering questions about the measures themselves, their administration, analysis, and use. After facilitators were assured that participants' questions had been answered group discussion focused on, (a) the importance of the measures, (b) their overall quality and comprehensiveness, (c) the feasibility of provider agencies using the measures, and (d) the utility of the performance measures developed and how providers could foresee using them.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

3.1 Contributions Towards Closing Care Gaps

This domain is optional for Spring 2025

4.1a Data Structure and Availability

The composite score from the Meaningful Activity measure will be available through electronic platforms, such as Qualtrics or incorporated into existing electronic systems providers already use. It is important to understand that HCBS data are by definition not medical data and therefore require different electronic systems from the usual hospital managed system. The HCBS field is not as centralized at the provider level and their electronic systems are developing.

That said, in addition to working with the systems providers may already been using, we are in the process of developing an electronic system at the University of Minnesota's Institute on Community Integration that houses this initiative. This system will offer providers an opportunity to utilize our electronic system and technical assistance to use the system to house their scores for a negotiated fee.

4.1b Implementation Costs and Burden

The measure score is derived from information from participants with disabilities or their proxies via an in-person or a Zoom interview. The cost and burden would be associated with the time it takes data collectors/someone designated to conduct these interviews to collect this data and how much they would be paid per interview/per hour. The measure itself would be calculated automatically using a formula embedded into an electronic data collection system that would also be used for performance reporting.

4.1c Confidentiality

As part of one of the nation's prominent research universities with a large academic health sciences program, the Institute on Community Integration and RTC/OM have access to a wide variety of servers/data storage systems that meet or exceed all HIPAA security requirements and can be used to ensure the privacy/confidentiality of personally identified information. All clinical and human subjects' data collected as part of RTC/OM performance measures will be secured with a University of Minnesota-approved resource at all times for the full extent of their life. This is true even when data has been de-identified. University approved methods of storing, clinical and human subjects' data that will be used by the implementer (RTC/OM) include servers exclusively devoted to and supported by the University's Health Sciences Technology (HST) group. HST supports University departments that need to store Private Highly Restricted data. The HST Operations and Infrastructure team supports and manages applications and infrastructure to meet University standards regarding HIPAA compliance. HST will work with RTC/OM staff to identify onboarding and maintenance requirements necessary to comply with UIS standards for PHR systems and data. Operating systems include professionally managed RHEL (Linux) and Windows Server environments. HST offers multiple storage options for data, documents, and folders to meet a variety of clinical, research, and storage needs. Network storage (SMB/NFS) is available on the Twin Cities campus for Healthcare Component (HCC) departments. The University's Health Information Privacy & Compliance Office (HIPCO) and its Institutional Review Board RB strongly encourage the use of limited datasets to maintain confidentiality when some identifiers are needed and this recommendation will be followed. A Limited Data Set is a dataset that contains a limited set of indirect identifiers, and it's only used within the University's Health Care Components or under a Data Use Agreement (DUA). This will allow implementers (RTC/OM) to access data with limited privacy risks while still being able to provide critical service quality and outcome data to RTC/OM measure users in a way that has a high degree of utility. Key aspects of creating a Limited Data Set will entail:

(a) *Removing Direct Identifiers:* All direct identifiers including names, street addresses, and telephone numbers will be removed.

(b) *Including Limited Indirect Identifiers:* Certain indirect identifiers, such as dates (e.g., date of birth, death, admission), and geographic information (city, state, zip code) will be included

(c) *Creating Data Use Agreements (DUAs):* When PHI in a Limited Data Set is shared with a third party, a Data Use Agreement (DUA) will first be established and approved by the UMN's IRB, Sponsored Projects Administration and Office of General Counsel. This agreement will outline the

permissible uses of the data and ensures compliance with HIPAA.

(d) *Review and Approval Processes:* The process for creating and sharing the Limited Data Set will include review by HIPCO, the IRB, and the Sponsored Projects Administration (SPA).

4.3 Feasibility Informed Final Measure

The final measure is a result of multiple feasibility assessments. First, we conducted Participatory Planning and Decision Making (PPDM) groups with people with disabilities, family members, and staff to weigh importance of the NQF domains and subdomains. This process assisted our team with the selection of measure concepts to prioritize for developing within HCBS. Second, the items and their response options from which the measure was composed was reviewed by a technical expert panel to test whether the measure reflected the intent behind the measure concept. Third, the measure underwent cognitive testing during which the items and their response options from which the measure was composed were tested by people with intellectual and developmental disabilities, aging needs, TBI, mental health needs, and physical disabilities. We wanted to assure that the measure was reflective of its measure concept as understood by people with different challenges. Fourth, the measure underwent pilot testing in two states (MN, PA). Feasibility was one of the main objectives of the pilot study during which we were able to identify items and response options that were either not providing us with information that reflected the measure concept accurately or those that did not contribute meaningfully to the measure. Based on all these stages to the feasibility assessment, we used an iterative process to refine the measure. The final version was used for psychometric testing in the RTC/OM National Field Study.

4.4 Proprietary Information

Not a proprietary measure and no proprietary components

4.4a Fees, Licensing, or Other Requirements

The measure is not proprietary, but the training and technical assistance will have an associated cost.

If organizations and users have the requisite knowledge (e.g., a Quality Assurance staff person with knowledge of HCBS processes) to use and score the measure then proprietary training would not be necessary.

Training that is available (online or in-person), focuses on providing potential users with background in order to do high quality measurement. More specific training on the measures we have developed, which is focused on appropriate use, administration and interviewing techniques, strategies for data analysis, and interpretation.

5.1.1 Data Used for Testing

The data presented in this submission were collected on a rolling basis during a multi-year field study of the RTC/OM measures between Spring 2021 and Spring 2024. This was a longitudinal data collection effort with three waves of data collection for each participant. However, only results from the first wave of data collection for each participant are presented unless otherwise noted. The first point of data collection for participants occurred between May 2021 and February 2024.

5.1.1a Dates of Testing Data

Field not required for Spring 2025.

5.1.2 Differences in Data

To mitigate autocorrelation and other statistical artifacts during data analysis only the first wave of data collection is used for data analysis, with one exception. The one exception is the test-retest analysis which could have been collected during any wave. Test-retest responses were only collected once per participant.

5.1.3 Characteristics of Measured Entities

Providers in the sample were recruited through a national directory of HCBS providers maintained by Medicaid.gov, as well as through networks of known HCBS providers and contacts recruited by University Centers for Excellence in Developmental Disabilities and other organizations contracted to support recruitment and data collection in several states (e.g., Utah, Pennsylvania, Florida, Georgia, Kansas). Additional HCBS providers were referred by participants who responded to national recruitment efforts (e.g., website, newsletters). Providers were not recruited for the study in every state, attempts to expand the representation of the sample were made in every state. This resulted in 67 organizations formally participating in the study across the states of Minnesota, Kansas, Florida, New Jersey, Pennsylvania, Massachusetts, Georgia, Arizona, Kentucky, Iowa, California, and New York. The size of participating organizations, in terms of the number of beneficiaries served, ranged from 10 or fewer to several hundred, with a variety of sizes in between. In Kansas, three large Managed Care Organizations also participated, as HCBS in that state are administered through them. The types of HCBS provided included residential services, in-home supports, home health or skilled nursing, employment services, community access, financial assistance, transportation, and more.

5.1.4 Characteristics of Units of the Eligible Population

HCBS beneficiaries were recruited either directly by participating provider organizations through direct outreach or recruitment materials (e.g., flyers, videos), or through national-level postings and newsletters inviting participation. Each participant was screened to verify eligibility, including age and receipt of HCBS or HCBS-like services. All participants in the study who

expressed interest in participating and met inclusionary criteria (age, currently receiving HCBS or HCBS-like services, able to provide consent/assent). Participant ability to understand the measure questions was first evaluated with the University of California, San Diego Brief Assessment of Capacity to Consent (UBACC). Capacity was also closely monitored by interviewers and if significant concerns about the validity of responses were raised, the participant's data were excluded.

Participants reported their primary disability as Intellectual or Developmental Disability (181, 61.4%), Physical Disability (59, 20%), Traumatic Brain Injury (24, 8.1%), Psychiatric Disability (17, 5.8%), Age-related Disability (4, 1.4%), or Other (10, 3.4%). The age range of participants was between 19 and 76 years old. Participants between ages 18-34 made up 37.9% of the sample, 35-54 were 44.7% of the sample, and the remaining 17.4% of the sample were 55 or older. 160 participants (54.2%) identified as male, 133 participants (45.1%) identified as female, and 2 participants (<1%) identified as "other". Participants identified their race as White (188, 63.7%), Black or African-American (66, 22.4%), Hispanic/Latino (12, 4.1%), Asian (2, <1%), or "Other race not listed" (5, 1.7%). No participants identified as solely American Indian or Alaska Native. Approximately 7.5% of participants identified with more than one race.

5.2.1 Level(s) of Reliability Testing Conducted

Person or encounter level (i.e., data element) (e.g., inter-abstractor reliability)

5.2.2 Method(s) of Reliability Testing

Internal consistency reliability and test-retest reliability methods were used to assess the reliability of person-level outcomes. Internal consistency reliability is a way to test the generalizability of a set of items to the broad domain of items that could have been used on the test. This type of reliability is used to gauge the level of error in content sampling of the items as well as errors of measurement arising from sampling, administration, or other secular effects. On the other hand, test-retest reliability is an estimate of errors around an examinee's "true" score over a short time frame (Crocker & Algina, 2008).

Internal consistency reliability of measure responses was assessed with Cronbach's alpha, a well-researched and widely-used method that was listed by the NQF as a method to demonstrate scientific acceptability (NQF, 2021). Test-retest reliability was evaluated using PPM correlations between participant composite scores across time points. Test-retest data were collected 10-14 days apart.

All calculations were performed in R (2025; version 4.4.0 or later) with the psych package (Revelle, 2025). A full data matrix of all respondents was loaded into R containing participant responses to all items on the instrument. For Cronbach's alpha calculations, a correlation matrix from these responses was computed using matrix smoothing and full information maximum likelihood (FIML), the latter an optimal technique for handling missing data (Enders, 2010). The sample size used was the average number of complete responses across all items. For test-retest reliability, missing data were handled by pairwise deletion.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

5.2.3 Reliability Testing Results

The attachment in 5.2.3a provides reliability testing results at the measure level. Standardized Cronbach's alpha and test-retest correlation statistics are highlighted in yellow. The detailed report on internal consistency also includes 95% confidence intervals as well as leave-one-out item analysis. Alongside the test-retest correlation coefficient is also reported the number of subjects and items in the sample.

5.2.3a Attach Additional Reliability Testing Results

[Reliability - Meaningful Activity.pdf](#)

5.2.4 Interpretation of Reliability Results

Internal consistency reliability (.92) and test-retest reliability (.94) were both excellent. These results demonstrate that scores obtained from the measure have a low amount of error and provide strong evidence that the instrument will consistently measure level of Meaningful Activity for participants.

5.3.1 Level(s) of Validity Testing Conducted

[Person or encounter level \(i.e., data element\) \(e.g., sensitivity and specificity\)](#)

5.3.2 Type of Accountable Entity Level Validity Testing Conducted

Not applicable/accountable entity level validity testing not conducted

5.3.3 Method(s) of Validity Testing

There were two approaches used to perform validity testing at the encounter-level. Content validity and construct validity were the two approaches used. An in-depth discussion of the content validation process undertaken in Study 1 of the RTC/OM project was provided in section 2.6 and will not be repeated here.

Construct validity of measure outcomes was evaluated with parallel analysis (PA) via scree plots as well as exploratory factor analysis (EFA). Both were used to evaluate the factor structure of each measure. Similar to reliability analyses, all calculations were performed in R. First, to determine the number of factors to retain during EFA, parallel analyses were performed then compared with the theoretical structure proposed during measure development. This helped guide the number of factors that were fit during the EFA procedure.

EFA models were fit with the psych package using ordinary least squares. Oblique solutions were produced with Oblimin rotation. Missing data were handled via full information maximum likelihood (FIML) when computing the correlation from the full data matrix (see section 5.2.2).

5.3.4 Validity Testing Results

The attachment in 5.3.4a provides the results of parallel analysis and exploratory factor analysis

for the Meaningful Activity measure outcome. Parallel analysis suggested retaining six factors, while the scree plot also indicates that there is a single strong factor which is in line with our initial one-factor hypothesis. The retention of six factors reflects the sub-groupings of items on the instrument, where six sets of four questions were asked based on six different life areas of a respondent. This result is in line with a bifactor interpretation of the data where there is a single strong factor and six associated subfactors for each of the four-item sets. This bifactor model was then fit and the results are presented below the parallel analysis results.

5.3.4a Attach Additional Validity Testing Results

[Validity - Meaningful Activity.pdf](#)

5.3.5 Interpretation of Validity Results

Parallel analysis suggested retaining six factors for the EFA model, which matched our hypothesized model of a 7-factor bi-factor model. Presented results show good fit statistics: TFI = .79 and RMSEA = 0.09. This model was also a much better fit to the data than the alternative one-factor model, i.e., one overall factor of Meaningful Activity. The six additional factors corresponding to the different areas of a person's life that are addressed by the measure shows this partitioning of factor variance to be a better fit to the data.

5.4.1 Methods Used to Address Risk Factors

Stratification by risk factor category

5.4.2 Conceptual Model Rationale

The National Quality Forum (NQF) emphasizes the importance of risk adjustment in evaluating outcome measures to ensure that potential threats to validity are addressed. Sociodemographic factors such as income and race have been explored as possible elements for risk adjustment by the NQF, which aims to develop guidelines in this area. In 2017, the NQF reviewed 303 submitted measures to assess their applicability for adjusting social risk factors that could affect health outcomes. The NQF panel recommended that these social risk factors follow the same criteria as clinical and health-related risk factors, although it noted a lack of a conceptual framework for their inclusion (National Quality Forum, 2014).

In the context of risk adjustment for this study, the NQF panel advised that sociodemographic factors should: (1) have a conceptual link to the outcome, (2) show an empirical relationship to the outcome, (3) display variability, (4) exist prior to intervention or care, (5) remain unaffected by intervention or policy changes, (6) be resistant to change, (7) be based on data that can be easily collected, (8) uniquely explain variations in the outcome, (9) contribute to the overall model, and (10) be considered valid and acceptable (NQF, 2014). These guidelines help differentiate risk adjusters from other variables.

More recently, the NQF conducted a review and convened a technical expert panel to develop further guidance for developers of outcome measures (National Quality Forum, 2020). They found that social risk factors mostly emerged at the individual and community levels, derived from various socioeconomic and demographic indicators. Functional risk factors, however, were often specific to individuals and based on self-reported survey data, with fewer clear definitions available. Statistical methods such as regression analyses were frequently used, though other models like hierarchical linear modeling were also applied to accommodate a broader range of risk factors.

Other risk adjustment models reflect similar themes to the NQF's recommendations. For example, the Centers for Medicare and Medicaid Services (CMS) sought expert input on risk adjustment, and the Department of Health and Human Services (HHS) reviewed and integrated those recommendations into 10 key principles for risk adjustment (Centers for Medicare and Medicaid Services, 2016). Five of these principles overlap with the NQF's guidelines, including recommendations that risk adjusters should be clinically relevant, predictive of medical costs, based on adequate sample sizes, encourage specific coding, and maintain internal consistency.

Additional support for the NQF's guidelines comes from the Research Agency for Healthcare Research and Quality (Velentgas, n.d.), which recommended that risk adjustment should not include variables affected by the outcomes, that variable selection should be based on prior knowledge of their relationship to outcomes, and that risk adjusters should have statistical ties to outcomes.

In our systematic review of studies involving risk adjustment for individuals with disabilities receiving home or community-based services, we categorized studies based on the type of risk adjusters used and their relation to specific outcomes. Panels helped prioritize the risk adjusters, and our findings suggest that four factors—chronic conditions, functional disability, mental health status, and cognitive functioning—may be recommended as candidate risk adjusters (Houseworth et al., 2022).

We have not currently collected data related to mental health status and cognitive functioning due to feasibility issues during the pilot study. We did collect data allowing us to stratify by functional disability/chronic conditions. Therefore, we are currently unable to determine if large difference between providers on those factors would impact performance scores, as the literature suggest. This could lead to some inappropriate conclusions.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

5.4.2a Attach Conceptual Model

[StratModeMA.png](#)

5.4.3 Variable Distribution Across Measured Entities

The attachment in 5.4.4a contains descriptive statistics across levels of functional disability for the measured outcomes. There are three levels of functional disability ranging from low to high service needs. Descriptive statistics reported for each level of functional disability are means, medians, standard deviations, minimum scores, and maximum scores.

5.4.4 Risk/Case-Mix Adjustment Modeling and/or Stratification Results

The attachment in 5.4.4a also contains an ANOVA analysis to determine if there were significant differences between the aforementioned functional disability categories on the measure outcome. No statistically-significant differences were found for the Meaningful Community-based Activity measure outcome.

5.4.4a Attach Risk/Case-mix Adjustment Modeling and/or Stratification Specifications

[Stratification - Meaningful Activity.pdf](#)

5.4.6 Interpretation of Risk/Case-mix Factor Findings

There were no significant differences between different levels of functional disability on whether or not a person was involved in activities that a person finds meaningful to them. Given that this is a new measure, we did not have expectations as to the direction of this relationship. Details on why other variables were not addressed in this model are explained in section 5.4.2

5.4.7 Final Approach to Address Risk Factors

Stratification by risk factor category

6.1.1 Current Status

Not in use

6.1.2 Current or Planned Use(s)

Quality Improvement with Benchmarking (external benchmarking to multiple organizations),
Quality Improvement (Internal to the specific organization)

6.2.1 Actions of Measured Entities to Improve Performance

The measured entities are bound by CMS regulations to provide HCBS services to support QoL outcomes in alignment with HCBS Final Settings Rule and Access Rule. As a minimum, this measure will provide evidence for the level of participation in meaningful activities, including engagement in activities at home and in the community to the extent they desire, expansion of natural supports and community connections, and service quality in supporting engagement in meaningful activities, of their clients as indicators of Access and HCBS. There is currently a gap in

measures available to provide reliable and valid data to service providers on HCBS outcomes, including the outcome of meaningful activity (UMN RTC/OM, 2020). The RTC/OM measure of Meaningful Activity has the potential to fill this gap for providers to be able to report to CMS on their progress toward HCBS and Access. Most importantly, providers will be able to make informed judgments about the way people with disabilities receiving HCBS experience their services and their effectiveness to improve/maintain their engagement in meaningful activities.

In our work with PAVE in CA, the measure score will be used to validate a PAVE measure score related to participation in meaningful activities to assess the quality of services as related to the Quality Incentive Program (QIP) for CA disability service providers. A pilot study is currently under way to collect data using RTC/OM and PAVE measures.

(A complete reference list is provided as a supplemental attachment in section 7.1.)

6.2.5a Potential Unintended Consequences

As with the use of any measure to increase outcomes, the use of the RTCOM Meaningful Activity measure may have potential unintended consequences, particularly if performance on the measure is tied to incentive or value-based payment programs. Poor performance on the measure could lead to reduced resources or negatively influence future funding decisions for provider organizations. In response, entities (HCBS providers) may reallocate resources from other critical areas to focus narrowly on improving measured meaningful activity outcomes, potentially undermining other areas of service quality. This can include a tendency to over focus on the measured outcomes and decreasing focus on other important or related outcomes. Another potential consequence is that providers may be held accountable for outcomes that are sometimes outside of their control. Not all individuals served have access to transportation, staffing, or community programs that directly support meaningful activity, limiting the provider's ability to influence the measured outcome. Additionally, implementing and reporting on the measure may require significant resources, especially for smaller or under-resourced providers. This may result in a reduction of resources for other performance measurement or quality improvement initiatives. To reduce the risk of these unintended consequences, the measure should be accompanied by clear guidance on appropriate use, interpretation, and its role within organization-wide and broader quality improvement strategies. It should not be used in isolation for value-based funding decisions and careful thought by users and policy makers should be put into equitable use of the measure in such initiatives. Potential unintended consequences for the beneficiary include pressure to participate in activities that may not be meaningful or desired by the individual, in an effort to improve scores on the measure. Organizations should ensure that individual preferences, goals, and desirable activities remain focused on when providing activities and when interpreting and responding to meaningful activity data. If used carefully, the benefits of the RTCOM Meaningful Activity measure such as improved attention to personal fulfillment, engagement, and enhanced service quality should outweigh the potential unintended consequences.

7.1 Supplemental Attachment

[Battelle Add References Single list.docx](#)

Developer POC email

abery001@umn.edu

Measure Developer POC

Brian Abery
Institute on Community Integration
Minneapolis, MN
United States

The measure developer is different from the measure steward

No

Steward Address

Brian Abery
Minneapolis, MN
United States

Steward Organization

Institute on Community Integration

Steward Organization URL

<https://ici.umn.edu>

Steward POC email

abery001@umn.edu