

Scoring — Learn What Helps

Joe Gleason · Advance review draft 3.7 — October 2026

At a Glance

The central idea: Measure whether life became better—not whether people spent more time in the system—and use what people discover to improve the next encounter.

What you'll explore: How to distinguish good design, useful participation, received support, and meaningful outcomes. Learn during a conversation, invite feedback afterward, and examine what changes over time. Recognize solutions contributed by patients, participants, caregivers, volunteers, and frontline teams—not only organizational leaders.

Design and AI: Explore optional real-time assistance that can bring forward an unanswered question, relevant resource, or possible next step. Use AI as a supportive facilitator and preparation tool, not an unquestionable referee or a hidden staff evaluator.

Put it to work: Create a **Benefit–Burden–Safety Measurement Plan** and a **Shared Learning Record** that connect observations and ideas with responsible action, contributor credit, and review.

Evidence and evaluation: Examine quality-improvement frameworks, person-reported outcomes, caregiver research, and an experiment in AI-assisted communication. Retain the **Jack Scale** as a historical and proposed review tool—not a validated clinical assessment.

“Ultimately, the secret of quality is love.”

— Avedis Donabedian, MD, MPH, healthcare-quality researcher, quoted by Michigan Medicine, July 29, 2016.⁷⁴

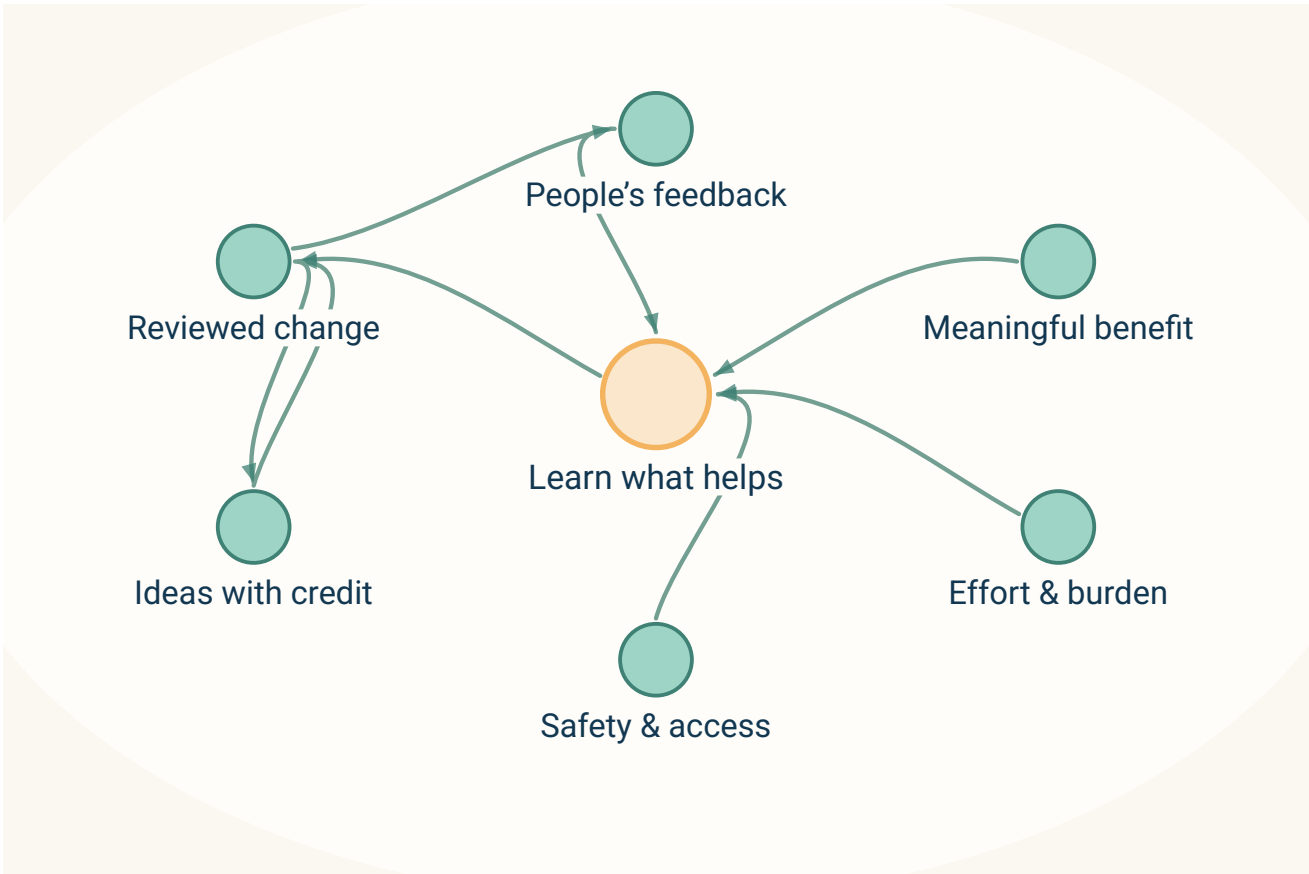


Figure 11 · Learning that gives credit

People’s feedback contributes to learning what helps. Examine meaningful benefit rather than treating system activity as the result. Examine the effort and burden alongside any apparent benefit. Examine safety and access, including people who were missed. The learning informs a reviewed change to try. Ideas inform the reviewed change with contributor credit attached. After a change, ask people again what helped and what should change. Tell contributors what happened to their ideas and let them choose how they are credited. A change is not automatically caused by the intervention. The historical/proposed Jack Scale is not a validated clinical assessment.

Notice and contribute	Review and return
People receiving and providing care	Ask what helped, what was difficult, and what should change. Preserve different accounts.
Responsible reviewer	Agree a proportionate change, check benefit and burden, and keep unsuccessful findings.
Contributors	Hear what happened to their suggestion; choose how they are credited. No hidden rankings.

In Context — What made the week better?

Toward the end of their review conversation, Sam asked Pat and Ellen a question that was not about the next appointment.

“What has been most useful?”

Ellen looked at the summary.

“Knowing what is actually arranged.”

“What difference does that make?”

“On a Thursday when everything happens as agreed, I can make a plan of my own. I’m not spending the morning checking whether someone is still coming or whether the day is confirmed.”

“And the messages?” Sam asked.

“Some are helpful. But more messages wouldn’t make Thursday more dependable.”

Pat had a different answer.

“The reading group. And the people at that table.”

“Is there anything you would change?”

“The activity afterward can get too busy. I’d rather do something quieter.”

Sam wrote that down separately.

“Does that mean you don’t want to go?”

“No. It means I like one thing and would change another.”

Ellen nodded.

“That’s how I feel about some of the information. I appreciate having it. I don’t need all of it every week.”

The service could be valuable while part of it needed adjustment. Information could be useful without more of it being better.

“What happens to what we’ve said?” Ellen asked.

“That is the important next question,” Sam replied.

They had offered more than feedback for a report.

They were helping design the support.

Scoring is a way to learn, not a verdict on people

The first nine principles have brought us from understanding a person to designing encounters, connecting pathways, improving access, and using AI with a defined

purpose.

The tenth asks whether that work helped—and how to make it better.

The original CarePhysics Scoring principle examines content, communication, access, coordination, and outcomes. It also calls for feedback to influence future design. This edition retains that purpose while distinguishing a reviewer's judgment, a person's experience, and evidence of an outcome.

Measure whether life became better, not whether people spent more time in the system.

Scoring does not require putting a number on every encounter. It means being clear about what we hoped to improve, what information could help us judge it, and what we will do with the answer.

Some measures are counts: requests answered, services received, or minutes required to complete a task. Others concern experience: whether someone felt heard, enjoyed an activity, or could use the time that support provided. Stronger claims about health, strain, quality of life, or financial value require an appropriate evaluation.

The person receiving care is not the object being graded.

Nor are the family's love and the staff's commitment things to infer from their activity records.

We are evaluating the support.

Learning belongs to everyone involved

Learning should not travel in only one direction—from an organization to the people it serves.

A clinician may find a clearer explanation. A patient may identify why an instruction will not work at home. A participant may suggest a better time to offer an activity. A caregiver may reveal where an arrangement creates unnecessary work. A volunteer may discover a more welcoming way to introduce someone.

Each contribution can help the next encounter.

The Day Center source materials reflect this approach: participants, families, care teams, social workers, volunteers, pilots, and focus groups contribute to the questions and designs being developed. Their experience is not merely a reaction collected after the software is finished.

For CarePhysics, that creates three connected opportunities.

During the encounter, notice whether something needs clarification or whether a useful option has emerged.

Immediately afterward, invite the people involved to describe what helped, what was difficult, and what they would change.

Over time, examine whether the chosen changes improve the experience, reduce burden, or contribute to the intended outcomes.

AI can assist at each point. The people remain the sources of lived experience, the interpreters of important differences, and the owners of decisions.

The aim is not simply to learn faster.

It is to make more of what people know available to one another.

Name what the number actually represents

The word *engagement* can conceal several different events. Keep the distinctions visible rather than combining everything into one activity score. The research review informing this book separates reach, understanding, first action, adoption, useful participation, outcomes, burden, equity, and safety.

What we are examining	The question	What would not establish it
Design quality	Is the material accurate, relevant, accessible, and ready to try?	A polished appearance or favorable internal rating alone
Reach	Did the intended person have a reasonable opportunity to encounter it?	Counting messages prepared rather than delivered
Understanding	Can the person explain the relevant idea, choice, or next step?	Opening an article or finishing a video
First meaningful action	Did the person take a useful step they chose?	Registration with no useful task completed
Adoption	Did they begin using the relevant service or support?	A request that has not led to use
Ongoing useful participation	Does participation remain worthwhile and appropriate?	Repeated logins, attendance, or a maintained streak alone
Experience	Did the person feel respected, heard, and able to influence what happened?	Staff or AI assuming the experience was positive
Outcome	Did something important change, such as strain, function, confidence, or quality of life?	A completed process without examining the intended change

What we are examining	The question	What would not establish it
Attributable impact	How much of that change can credibly be attributed to the intervention?	A favorable before-and-after difference without considering other explanations

These are not compulsory stages everyone must complete.

A person may understand an option and reasonably decline it. A caregiver may stop using a tool because the immediate need has been resolved.

For Pat's family, the reading guide was offered. They chose to try it. They described what they enjoyed and what was confusing.

That gives the team information about use and experience. It does not establish improved cognition or a benefit caused by the technology.

The distinction does not diminish a pleasant evening. It keeps the claim faithful to what happened.

Measure progress through the pathway—not speed alone

A constraint-focused project needs to know whether the whole pathway improved. Keep operational measures alongside outcomes and balancing measures that examine unintended consequences. IHI recommends this combination; it is not a new CarePhysics score.⁷⁵

Record time to the first substantive response, time to a confirmed arrangement, and time to the service itself as different intervals. These are lead-time measures. A service delivery rate counts defined services actually delivered per week or another stated period. It is not the same quantity as days of waiting, and it is not TOC's financial throughput-accounting measure.

Keep unresolved demand visible: how many requests remain open, how long they have been open, their stage, known reason, and responsible responder. Report completed-case waiting times alongside the age of open requests. Otherwise, those waiting longest disappear from the apparent improvement. An automatic acknowledgment is not necessarily a substantive answer.

Choose units and denominators before comparing results. Preserve clinically appropriate priority, count missing outcomes, and examine language and access differences without guessing identities. Do not improve a queue statistic by deleting requests, moving work to families, or selecting only easier cases. Necessary conversations should not be rushed to meet a rate target. See Appendix F6 for local definitions and an arithmetic example.

Learn while the conversation can still change

A question does not have to wait for the next evaluation report to be answered.

With appropriate agreement, a configured assistant could help identify an unanswered question, an unfamiliar term someone has asked about, or a proposed next step without an accepted owner. It could bring forward a relevant approved resource or suggest a clarification for the facilitator to consider.

The dated Genus Coach description includes consented real-time transcription and timely suggestions, while explicitly making family-facing use dependent on deployment and configuration. That is a documented capability description—not evidence that every conversation is accurately assessed or improved.

A useful suggestion might read:

Possible clarification: Ellen asked who will contact her if the arrangement changes. That responsibility has not yet been confirmed in this conversation.

Or:

Relevant resource: The center’s approved guide explains the introductory options being discussed. Check that the version and arrangements are current before using it.

The facilitator decides whether to use the suggestion. It should be easy to dismiss, and it should not interrupt every pause.

This allows us to examine specific parts of a conversation while it is happening. Was a question raised? Was an answer offered? Was an action accepted? Does the person say the explanation helped?

Those are different from an assistant declaring that someone is reassured, confused, resistant, or trusting.

Do not treat facial expression, tone, silence, or speaking time as an authoritative measure of private feelings or comprehension. In this proposed approach, explicit statements and confirmed actions carry more weight than inferred emotional labels.

A transcript can also contain errors. People must be able to correct the record.

Real-time assistance can identify a moment worth checking. The people involved establish what that moment means.

Bring a possible solution forward without taking credit for it

In Context — The useful idea comes from Pat

In this proposed version of the review, everyone has agreed to limited live assistance after Sam explains what it processes, who sees its suggestions, and how it can be paused. The meeting can proceed without it.

Sam and Lena were discussing Pat’s request for a quieter afternoon option.

“We can review the activity schedule,” Lena said.

Pat looked at her.

In Context — The useful idea comes from Pat

“Could you tell me the choices before lunch?”

“Would that help?”

“Yes. Sometimes people ask me what I want when I’ve already had enough of deciding.”

The assistant brought forward a brief reminder for Sam:

Pat’s proposal: Explain the actual afternoon choices earlier. Confirm the preference before moving to other arrangements.

Sam did not describe it as the assistant’s discovery.

“Pat, you’re suggesting that the timing matters—not only which activities are available.”

“Exactly.”

The assistant also located the center’s reviewed activity-options guide. Lena checked what it said against the arrangements her team could actually provide.

“We could try discussing the available choices earlier on the days you attend,” she said. “I need to check with the team how we would do that consistently.”

“And I can still change my mind?”

“Yes.”

“Then we have a promising proposal.”

The assistant could help record the suggestion and prepare the follow-up. It could not promise staffing, invent an activity, or decide that the change had worked before anyone tried it.

The useful idea came from Pat.

Lena accepted responsibility for checking feasibility.

A later conversation would establish what happened.

Ask everyone what helped

After an encounter, offer a brief opportunity to reflect while the experience is still fresh.

A patient, participant, caregiver, clinician, staff member, or volunteer may notice something the others missed.

A proposed invitation could say:

Help us improve the next conversation

You can tell us what helped, what was difficult, or what remained unanswered. Positive and critical responses are equally useful.

What should we keep?

What should we change or try?

You may answer now, respond later, speak with a person, or skip the questions. Before you respond, we will explain who receives your feedback and how to request a reply.

The organization must supply the actual recipient, response arrangements, and privacy information. These are local improvement questions, not a validated questionnaire.

An assistant could help someone organize a spoken or written answer, where that format is supported and agreed. The person should be able to check what will be shared and change wording that does not sound like them.

Do not automatically send every response to everyone who attended.

A caregiver may need a private route. A volunteer may want help reflecting on their own communication. A patient may prefer to raise a concern with someone other than the professional involved.

Staff can also be offered private, optional practice: “Help me explain this more clearly next time.” That should not quietly become a manager’s performance report.

Keep differing accounts visible. “The clinician believed the plan was clear” and “the patient remained unsure whom to call” should lead to a better explanation, not a contest over whose rating is correct.

A supportive third party, not an unquestionable referee

AI may offer another place to prepare a concern or suggest an improvement. Whether people find that more comfortable than speaking directly is something to ask—not assume.

The appropriate aspiration is **a nonjudgmental facilitator designed to treat perspectives fairly**, not an inherently neutral judge.

AI systems can produce inaccurate, biased, or incomplete statements and encourage inappropriate reliance. WHO’s guidance emphasizes defined tasks, stakeholder participation, transparent responsibility, and continuing assessment.¹

For this application, the assistant should neither defend the organization nor automatically agree with the speaker.

It could help a caregiver say:

“I appreciate the support. I still do not know who will call if the plan changes.”

It should not convert that concern into “The organization is failing you” or “You should

be grateful.”

The service should explain who configures the assistant and who can access the feedback. An organization-operated channel must not present itself as an independent complaints authority.

When conflict, safety, or professional judgment requires a person, provide an appropriate human route. Agreement among several relatives does not make Pat’s preference disappear. A favorable average does not cancel a serious concern.

Fairness requires a process people can question.

Capture wins with context and credit

A learning system should notice what went well as well as what failed.

A participant finds a comfortable way to contribute. A patient suggests a clearer preparation checklist. A caregiver identifies a better time for an update. A volunteer demonstrates how to use a shared screen without taking over. A staff member finds words that make a difficult explanation easier to discuss.

These moments can be recognized without claiming more than they establish.

“Capture a win” should mean preserving what was useful, who contributed, and the conditions under which it helped—not collecting only flattering stories.

The dated Day Center materials describe Team Pulse as a place for wins and observations, alongside family feedback and editable updates. The description supports a contribution-and-review workflow; it does not establish that every contribution leads to improvement.

A Shared Learning Record can be short:

What needed attention: Pat wanted a quieter option after the reading group.

Who contributed the idea: Pat suggested discussing choices earlier.

What will be tried: Lena checks a workable arrangement with the team and Pat.

What will be reviewed: Pat’s experience, whether the choices were genuinely available, and the work required.

What may be shared: A general lesson, with permission and without unnecessary private details.

The record develops as the work happens. It should distinguish an idea from a trial, a reported benefit from a repeated finding, and a local result from evidence that applies elsewhere.

Contributor credit is itself a choice. Ask how someone wants to be acknowledged. A participant should not have to disclose a diagnosis or personal story to receive recognition.

And keep the result that did not improve. Otherwise the collection becomes promotion rather than learning.

Let the learning change the design

Feedback should be able to change a video, article, survey, message, conversation, or pathway.

The reading guide already offers an example. Ellen initially thought the optional prompts were instructions. That observation could lead the team to separate suggestions from the story and make their optional nature more visible.

Pat's suggestion concerns timing. The next improvement may be a conversation earlier in the day rather than a new piece of content.

In a separate illustrative example, a volunteer notices that families cannot easily tell which door to use. The answer might be an accurate arrival photograph, a revised message, or someone meeting visitors—not a longer welcome booklet.

Use the familiar design functions: purpose, main idea, demonstration, chosen next step, questions or connection, and what follows. Preserve recognizable structure while adapting the language, examples, format, and assistance.

Ask how culture, beliefs, resources, and local circumstances affect the experience. Do not assume one community's response identifies the right design for every other community.

Clinical instructions and validated questionnaires retain their required meaning, wording, and scoring. Clarity can be added around them without casually rewriting the instrument.

The learning cycle is practical:

Notice → ask → consider options → try a chosen change → hear from those involved → review → share appropriately → check again.

Each transition needs responsibility.

Feedback submitted is not feedback considered. A proposed solution is not a change implemented. A revised article is not evidence that the next reader understood it.

The Jack Scale: preserve the question behind the number

The **Jack Scale** belongs to the history of CarePhysics.

The framework's history credits Jack Gleason with proposing a scale from **-5 to +5**, with **zero as neutral**, to recognize that an experience may detract from its purpose as well as contribute to it.

Its underlying question remains useful:

Did this experience help, do little, or make things harder?

The Jack Scale remains a historical and proposed review tool. Clinical validation or reliable outcome prediction has not been established in the sources reviewed for this book.

An organization choosing to use it should define the specific item and what its ratings mean. Keep the reason and evidence beside the number.

“Reviewer judged the next step unclear” is useful.

“Family engagement is -2” conceals too much.

A team can record what it expects a redesigned item to improve, then compare that expectation with actual experience. Do not revise the original prediction afterward to make the design appear more successful.

Zero must not mean “we have no information.” Missing, declined, not applicable, and not yet reviewed need separate labels.

Avoid false precision. Neighboring ratings may be difficult to distinguish. Reviewer disagreement may reveal an unclear standard or differing assumptions about the audience.

A favorable total also cannot cancel a serious problem. An attractive video with an incorrect safety instruction is not ready because its other ratings are high.

Use the scale to support discussion where useful. Do not use it to rank families, infer staff compassion, determine eligibility, or claim clinical benefit.

Choose measures that can change a decision

Start with the difference the support is intended to make.

For Ellen, that may be dependable time she can use without continuing to coordinate care. For Pat, it may be a worthwhile day with choices. For staff, it may be an accurate update that does not require after-hours reconstruction.

IHI recommends examining **outcome measures**, **process measures**, and **balancing measures**. Balancing measures ask whether an improvement in one area creates problems elsewhere.⁷⁵

For live conversation assistance, a process measure might examine whether a question received a response. The experience measure asks whether the person found the response useful. A balancing measure examines distraction, review effort, or discomfort about being recorded.

For a family update, timely delivery matters. So do accuracy, understanding, and the work required to prepare it.

A shorter interaction is not necessarily better. A useful conversation may take longer. More questions may indicate confusion—or greater willingness to speak.

Interpret the measure through the purpose, not through a general preference for higher or lower numbers.

A local question can be sufficient for improving tomorrow’s activity. A claim of reduced caregiver strain requires a more systematic approach and appropriate measurement.

When using a standardized instrument, verify the construct, respondent, setting,

language version, permissions, recall period, scoring, and interpretation. An AI rewrite does not automatically preserve its measurement properties.

An intake baseline can help begin a conversation. Repeating a measure requires an agreed process and suitable interpretation; the initial score alone cannot establish change. Appendix J preserves the dated local-intake example.

Private caregiver concerns also need an appropriate confidential route. Do not collect sensitive answers in a shared family record and rely on wording alone to prevent disclosure.

An assessment needs a response and a return point

RUSH's Caring for Caregivers (C4C) model is one public example of assessing the caregiver's own needs and returning to them over time. Its published description names caregiver-burden, anxiety, depression, health-literacy, and social-needs measures. Those measures are not interchangeable with the local four-item intake described above.⁷⁶

For a local evaluation, name the instrument, version, language, respondent, baseline event, and follow-up dates. Enrollment, the first assessment, and completion of a support program are different starting points. Compare the same intended construct while recording any changes in the tool or delivery.

Our proposed process permits a person to decline a nonmandatory question or choose a private conversation. Record declined, not asked, and missing separately; none means zero strain. Follow the instrument's rules for incomplete responses rather than silently changing its score.

Before asking sensitive questions, establish the authorized reviewer, response time, private record, and professional escalation route. A concern requiring prompt attention must not wait for the next scheduled measurement. Appendix F turns these requirements into a planning card.

Keep missing people and different perspectives in view

Ask the person directly where they can express the relevant experience, with appropriate assistance. Label caregiver or other proxy accounts separately. Staff observations should remain observations.

A person can sit through an activity without feeling that they belong. A caregiver can appreciate a service and still find the arrangements exhausting.

Consider an **invented reporting example**: a center invites 20 households to provide feedback. Twelve respond, and nine say the guide was useful.

"Seventy-five percent of responding households found it useful" describes nine of twelve.

"Seventy-five percent of families benefited" does not.

Eight households have not provided feedback. The other three responses also need understanding. Silence is not satisfaction, and missing information is not zero.

Define the denominator—the group a percentage describes—and the unit being counted. People, households, encounters, requests, and service days are not interchangeable. Several answers from one household are not several independent households.

Offer appropriate languages, formats, times, and assistance. Include people who decline digital or AI participation. Do not infer cultural preferences or demographic characteristics from names, addresses, or communication style.

The export explicitly states that its described intake lacks demographic fields needed for some proposed group comparisons. That limits what can be concluded from those records alone.

Differences among communities may reflect resources, care needs, transport, service capacity, or access—not simply the quality of staff effort.

Distinguish a change from its cause

A family reporting less strain is important. It does not, by itself, establish what caused the change.

During the same period, relatives may have accepted responsibilities, clinical care may have changed, or the household may be going through a different stage of need.

Match the comparison to the claim.

To examine whether an evening note adds value, compare communication approaches while considering the underlying service. To examine live AI support, compare the relevant encounter process, including preparation, review, and follow-up—not an assisted meeting against an unrelated experience.

Deterioration does not automatically mean that care failed. Comfort, choice, practical support, and family capacity can matter while an illness progresses. Stability alone also cannot prove that a program prevented decline.

Use research expertise and an appropriate comparison when stronger causal claims are intended. Report uncertainty and results that are too imprecise to support a conclusion.

A familiar theory can explain why an approach is worth trying. It cannot supply the missing outcome evidence.

Nine outcome areas, not nine promised results

CarePhysics identifies nine outcome areas. Retain them as questions to examine, not a promise that every application improves all nine.

Original outcome area	A question for evaluation
Improved Culture of Care	Can people raise concerns, contribute ideas, and see respectful responses?

Original outcome area	A question for evaluation
Integration of SDOH	Do social needs, such as transport or affordability, lead to practical assistance?
Care Equity	Who can reach and benefit from support, and who remains excluded?
Community Connections	Do introductions become appropriate services or useful relationships?
Stronger Relationships	Are communication, commitments, and opportunities for disagreement improving?
Improved Quality of Life	Does something important in the person's life improve?
Reduced Burnout	What happens to staff workload and well-being—and separately to caregiver strain?
Improved Outcomes	Do relevant health, function, safety, or care outcomes change under an appropriate evaluation?
Improved Bottom Line	Is the service sustainable after implementation, staffing, review, and other costs?

Choose the areas relevant to the project. Do not administer every possible questionnaire merely to fill the table.

Keep financial perspectives separate. A saving for an organization may shift work or expense to a family. Better documentation may support a funding application without guaranteeing payment.

Keep **willingness to recommend**, **an actual referral**, and **a referred person receiving useful support** separate too.

Recognition should reward contribution, not favorable answers. Negative feedback deserves the same opportunity to influence the work. Points or participation totals must not become family leaderboards, care-quality ratings, or promised funding.

Build a Benefit–Burden–Safety Measurement Plan

Begin with one decision and one manageable process.

For example, an organization could test whether optional live assistance and brief post-meeting feedback help people leave a review conversation with their concerns represented and responsibilities clear.

Start with appropriately authorized practice cases. Then, where the organization approves a real-world test, establish a baseline and a defined review period. A two-week baseline followed by a four-week improvement test is an illustrative planning option—not a design sufficient to establish long-term clinical effectiveness.

Record the population, source materials, configuration, pathway version, reviewer, and authority to change or stop the process.

What to examine	Measure and source	Responsibility and possible action
Questions and responsibilities	In a reviewed sample, questions with an answer or accepted follow-up owner divided by questions identified. Record what remains unresolved.	The facilitator checks the record and repairs missed follow-through.
Experience and choice	Optional feedback on being heard, ability to disagree, usefulness of assistance, and preference for future encounters. Report invitations, responses, and declines.	The service lead reviews differences and changes the offer or format.
Ideas becoming improvements	Accepted improvement actions completed by their review date divided by those due; retain suggestions still awaiting a decision.	A named improvement owner records the action and returns an answer to the contributor.
Meaningful benefit	A suitable measure of the project's intended result, such as understanding, usable caregiver time, or participant experience.	The relevant professional interprets findings within their role and adjusts support.
Burden and access	Preparation, encounter, review, correction, and follow-up time; distraction and access difficulties across available routes.	The operational lead revises staffing, training, prompts, or workflow.
Safety and accuracy	Unsupported suggestions, misattributed statements, privacy errors, missed escalation, and incorrect service information.	The designated safety and service owners act promptly; serious concerns do not wait for a scheduled review.

For conversational signals, the assistant's count should not establish its own accuracy. Compare a sample with appropriate human review. Examine both misleading prompts and important matters the assistant missed.

Count the whole task. Faster drafting may be offset by checking and correction. A benefit to the team may shift work to the family.

Set review criteria in advance. A serious privacy failure is not balanced away by favorable satisfaction responses.

Let AI assist the learning without inventing it

An assistant could help organize permissioned feedback, identify missing definitions, retrieve relevant guidance, or prepare a draft improvement report.

Its assignment should make the distinctions explicit:

Using the approved definitions and authorized records, summarize what was offered, what people experienced, and what changed. Preserve disagreements and negative findings. Show denominators and missing information. Credit contributors according to their permissions. Separate proposed solutions, actions completed, and outcomes. Identify what cannot be established.

Use checked calculations and reconciled records for numerical reporting. Fluent prose does not verify arithmetic or data selection.

The dated Day Center record distinguishes booked from attended days, does not capture family-guide completion, and has no separate referral-date field on the adult day intake. An assistant must preserve those limits rather than generate the report someone hoped the records could support.

Reviewed research and design guidance can help the team develop possible responses. Current local information establishes which resources actually exist. Personal context shapes a response only within its permissions.

Keep those sources distinct, with owners, review dates, approved uses, and corrections. A correction may lead to a revised instruction, template, source entry, or workflow; it does not automatically retrain a model. The knowledge-bank companion remains a proposed way to organize this work.

Properly applied, AI can reduce repetitive assembly, capture useful experience, and give people more attention for interpretation and action.

It should not infer a hidden score of love, cooperation, kindness, or entitlement to help.

Give the organization the capacity to learn

A supportive learning process needs time, trust, and authority.

Explain what live assistance captures, what is retained, who sees suggestions, and which feedback remains private. Separate permission for the current encounter from permission to reuse a story or example. Allow people to decline or pause assistance and

retain appropriate human, telephone, paper, and accessible-language routes.

Do not promise anonymity when responses are identifiable. Explain any relevant limits on confidentiality and required safety reporting.

Practice with fictional or appropriately authorized material. Include mistakes and omissions in a training environment; do not seed false statements into real care records to test staff.

Provide time for review, technical support, interpretation, training, and responding to concerns. An invitation to contribute creates work for someone who must listen and act.

AI may make routine information or question preparation available around the clock. That is not continuous professional coverage. State actual human response hours and the appropriate urgent routes.

Let staff help shape the process. Use their ideas to improve workflow rather than interpreting every difficulty as a performance problem. Clinical accountability remains; undisclosed scoring of compassion is not a substitute for it.

The same principles apply in health systems, home care, independent living, and community programs. A patient may improve a procedure-preparation explanation; a volunteer may improve a welcome; a caregiver may identify a broken handoff. Their contributions need review appropriate to the task, not dismissal because they came from outside the professional team.

At community and state levels, share definitions and lessons without exposing private stories or ranking unlike populations. Relationships and better messages cannot replace unavailable services.

Genus supplies technology. Partners retain responsibility for care, staffing, programs, professional decisions, and follow-through.

Make the learning visible

Return to the contributor.

“You suggested this. We tried it. Here is what happened.”

Where the idea cannot be implemented, explain the reason and any alternative. Where it did not help, keep that result.

The dated export proposes a register containing the finding, decision, responsible person, actual change, date, and remeasurement. It explicitly says that this register is maintained outside the described product. A commitment to learn is not evidence that learning happened.

A learning record should therefore be able to show a decision to continue, adjust, narrow, stop, or make no change.

Sometimes the most useful improvement is removing a field, reducing a message, or restoring a conversation with a person.

The measure earns its place when it helps someone make a better decision.

Models and Evidence Behind This Chapter

The chapter combines the original measurement framework with the proposed real-time and post-engagement learning approach. The sources below support different parts of that design; none validates the complete CarePhysics application.

Donabedian's framework and improvement measures

Examine conditions, processes, and results

Donabedian's 1966 paper, reprinted in 2005, distinguishes the structure of care, the processes through which it is provided, and outcomes. It also warns against assuming that the relationships among them are established. This is methodological work, not an intervention trial.⁷⁷

IHI's guidance complements that perspective with outcome, process, and balancing measures and repeated review over time.⁷⁵

Where we used it: A conversation completed, a question resolved, a caregiver's experience, and staff workload receive different measures.

Local question: Did the process produce the intended value without unacceptable consequences elsewhere?

AI-assisted conversational feedback

Assistance can be useful before a response is finished

Ashish Sharma and colleagues' 2023 randomized study tested **HAILEY** with **300 peer supporters** from TalkLife. The system offered just-in-time feedback while supporters prepared written replies. The authors reported a **19.6 percent increase on their measure of conversational empathy** overall.⁷⁸

Where we used it: The proposal to offer limited suggestions while a person still controls the response.

Evidence boundary: This was a nonclinical, text-based study—not a trial of spoken dementia-care meetings or Genus. A higher score for empathic language does not establish improved clinical outcomes, reliable emotion detection, or impartial mediation.

Local question: Does the assistance help people respond usefully while preserving their judgment and attention?

Person-reported measures

A questionnaire works through a relationship

Joanne Greenhalgh and colleagues' 2018 realist synthesis examined how patient-reported outcome measures influence communication and care. The wider synthesis included **46 papers**, with **42 informing the article**.

Measures could help people reflect and raise concerns, but could also constrain

conversation depending on their use and the response they received. The review explains mechanisms and context rather than providing one pooled effect size.⁷⁹

Where we used it: Immediate feedback leads into dialogue and action rather than replacing them.

Local question: Did asking the question give the person a useful voice?

Caregiver outcomes

Measure the support people are meant to receive

Steven Zarit and colleagues' Daily Stress and Health study followed **173 family caregivers of people with dementia over eight consecutive days**. Caregivers reported fewer care-related stressors and more positive experiences on adult-day-service days, alongside more noncare stressors. It was a within-person observational study, not randomized attendance.⁴³

Laura Gitlin and colleagues' ADS Plus trial involved **203 caregivers across 34 sites**, randomized by site to usual services or services with structured caregiver support. Adjusted depression scores were lower in the added-support group at 12 months. **22.7 percent** were lost to follow-up, the sample was predominantly female and college educated, and the attendance comparison did not meet the conventional statistical-significance threshold.⁴⁴

Where we used them: Caregiver experience and well-being are examined separately from service activity.

Evidence boundary: Neither study establishes an effect of AI assistance, a family summary, or this proposed measurement process.

Person-centered measurement in Day Centers

Presence is not the same as belonging

Clara Scher and colleagues' 2025 e-Delphi study involved **22 practitioners and researchers** reviewing candidate measures. The panel reached consensus on selected measures of meaning and purpose, friendship, and belonging, but not on a single engagement measure.

This was expert-consensus work, not real-world validation of every selected instrument or evidence that Day Centers improve those outcomes.⁴⁵

Where we used it: Pat's preferences and experience remain distinct from attendance.

Local question: Can the organization learn what matters without making measurement another burden?

Behavioral models, social learning, and responsible AI

Let findings change the response

COM-B helps investigate barriers. Self-Determination Theory directs attention to chosen

participation. Social Physics offers a perspective on how useful knowledge travels through relationships.

The research review informing this book distinguishes their evidence: strong conceptual support for COM-B, moderate intervention evidence for Self-Determination Theory, and emerging-to-moderate support for the broader Social Physics perspective. Those judgments do not transfer automatically to a scoring system or live assistant.

WHO's guidance adds requirements for defined tasks, stakeholder involvement, privacy, and protection against overreliance. It is governance guidance, not proof of effectiveness.¹

Local question: Does the learning process strengthen people's agency and the support around them—or merely increase what the organization can count?

One thing to try

Choose one encounter your organization offers regularly.

Ask a person receiving support and a person providing it what helped, what was difficult, and what they would change. Offer separate responses and appropriate privacy.

Select one contribution, name an owner, and try a manageable change. An approved assistant may help organize the information or prepare options; people decide what to use.

Then return to the contributors and examine what happened.

Alongside that exercise, take one success measure you already report and ask:

What important human outcome could become worse while this number improves?

Add the question or measure that would help you see it.

Ten principles become an ongoing practice

In Context — “Ask me again”

Before the meeting ended, Sam read back the actions—not an overall score.

Lena would check how to discuss afternoon choices earlier with Pat. The team would continue reviewing the reading guide so optional prompts looked optional. Ellen's support questions would remain part of the ongoing conversation, with private concerns handled through the appropriate route.

The additional Tuesday had no new confirmed resolution. It remained an open need.

“How will we know whether the afternoon change helped?” Sam asked Pat.

In Context — “Ask me again”

“Ask me again.”

“And if it doesn’t?”

“I’ll probably mention it.”

Ellen smiled.

“That is dependable.”

Lena asked whether Pat was comfortable with the team using his suggestion as a general example.

“As long as you say it’s a suggestion,” he replied. “I haven’t inspected the results.”

They had not solved every problem.

They had made one improvement specific enough to try, preserved the name of the person who proposed it, and agreed to find out whether it helped.

The tenth principle turns CarePhysics into an ongoing practice of listening, contributing, trying, checking, and changing. The next five chapters apply these ten principles to the people and organizations doing the caring.

Notes

1. World Health Organization (2024). WHO releases AI ethics and governance guidance for large multi-modal models. January 18. [Source](#)

[Return to text 1](#) · [Return to text 2](#)

43. Zarit SH, Kim K, Femia EE, Almeida DM, Klein LC (2014). The effects of adult day services on family caregivers’ daily stress, affect, and health: outcomes from the Daily Stress and Health (DaSH) study. *The Gerontologist*. 54(4):570–579. DOI: 10.1093/geront/gnt045. [Source](#)

[Return to text](#)

44. Gitlin LN, Roth DL, Marx KA, et al. (2024). Embedding Caregiver Support Within Adult Day Services: Outcomes of a Multisite Trial. *The Gerontologist*. 64(4):gnad107. DOI: 10.1093/geront/gnad107. [Source](#)

[Return to text](#)

45. Scher CJ, Anderson K, Zagorski W, Siamdoust S, Finik J, Sadarangani T (2025). Identifying Person-Centered Outcome Measures for Use in Adult Day Services: An E-Delphi Consensus Study. *Sage Open Aging*. DOI: 10.1177/30495334251408570. [Source](#)

[Return to text](#)

74. Michigan Medicine (2016). How Love and a Landmark Paper Improved Health Care. July 29. Historical account quoting Avedis Donabedian. [Source](#)

[Return to text](#)

75. Institute for Healthcare Improvement. Model for Improvement: Establishing Measures. Guidance on outcome, process, and balancing measures. [Source](#)

[Return to text 1](#) · [Return to text 2](#) · [Return to text 3](#)

76. Carbonell E, Mariani D, Golden R (2025). Caring for Caregivers Within Age-Friendly Health Systems: An Organizational Case Study of National Scale and Spread. *INQUIRY*. 62:00469580251325660. DOI: 10.1177/00469580251325660. [Source](#)

[Return to text](#)

77. Donabedian A (2005; originally published 1966). Evaluating the quality of medical care. *The Milbank Quarterly*. 83(4):691–729. DOI: 10.1111/j.1468-0009.2005.00397.x. [Source](#)

[Return to text](#)

78. Sharma A, Rushton K, Lin IW, et al. (2023). Human–AI collaboration enables more empathic conversations in text-based peer-to-peer mental health support. *Nature Machine Intelligence*. 5:46–57. DOI: 10.1038/s42256-022-00593-2. [Source](#)

[Return to text](#)

79. Greenhalgh J, Gooding K, Gibbons E, et al. (2018). How do patient reported outcome measures (PROMs) support clinician-patient communication and patient care? A realist synthesis. *Journal of Patient-Reported Outcomes*. 2:42. DOI: 10.1186/s41687-018-0061-6. [Source](#)

[Return to text](#)