ARTICLE

Caregiver-Assessed Outcome Measures Demonstrate Multi-Domain Improvements in Sleep, Focus, Mood, Energy, and Memory with 40 Hz Invisible Spectral Flicker Therapy: A Pilot Study at Masonicare (n=10)

Pedro Vieira[1,2], Paula A. Smith[3], Lauren Dubuque[4], Anne Collette[5], Jesse Gomez[6,*], Mai Nguyen[7], Marcus Carstensen[8]

[1] Department of Health Technology, Denmark Technical University, Copenhagen, Denmark | [2] Data Analyst, Optoceutics Inc., 2120 University Ave, Berkeley, CA 94704 | [3] Memory Care Resource Coordinator, Masonicare at Mystic, 45 Clara Drive, Mystic, CT 06355 | [4] Executive Director, Masonicare at Mystic | [5] Chief Strategy Officer, Masonicare at Mystic | [6] Chief Revenue Officer, Optoceutics | [7] Chief Executive Officer, Optoceutics | [8] Chief Scientific Officer, Optoceutics | [*] *Corresponding authors: Jesse Gomez, Chief Revenue Officer* | jesse@optoceutics.com | optoceutics.com

---

Background: Dementia care imposes a profound and growing burden on residents, caregivers, families, and health systems. 40 Hz Invisible Spectral Flicker (ISF) therapy represents a novel non-invasive assistive technology with established preclinical and clinical evidence for cognitive support. While self-reported outcome measures have provided encouraging real-world evidence in home-user populations, caregiver-assessed observations offer complementary advantages in long-term care settings, where cognitive impairment limits self-report reliability and where the therapy's impact on staff burden and care economics is as consequential as its effect on individual residents. **Methods**: We conducted a 6-week pilot study at Masonicare, a US-based long-term care facility, collecting caregiver-assessed outcome measures from 10 residents receiving 1 hour daily of 40 Hz Invisible Spectral Flicker therapy. Trained caregivers completed weekly assessments across five domains—Sleep, Focus, Mood, Energy, and Memory—using a categorical scale (Improved/Stable/Worse) converted to algebraic scores (+1/0/-1) for quantitative analysis. Only one participant withdrew, after week 2, and 9 of 10 residents completed all 6 weeks of assessment. **Results**: Caregivers observed consistent positive improvements across all five assessed domains throughout the 6-week period. Mood and Sleep exhibited the strongest net improvement rates (67.9 and 64.3%, given by the number of Improved minus the number of Worse assessments), followed by Focus (44.6%), Energy (28.6%), and Memory (25.0%). Notably, from week 3 onwards, caregivers reported zero "Worse" assessments across all domains, indicating sustained stabilization or improvement. The cumulative absolute improvement across all domains totaled +129 (sum of Improved minus Worse assessments), with an average cumulative improvement score of +2.98 (out of a maximum of 6) per resident by week 6. Caregivers also reported decreased or very decreased burden when caring for the residents after the 6 weeks, with family members noting improvements in residents' mood and engagement. All but one resident demonstrated very reduced burden for their caregiver across 4 domains, with +1 to +3 point improvements on a 5 point scale. **Conclusions**: This pilot study provides preliminary evidence supporting multi-domain improvements in resident quality of life, more independence in activities of daily life, with positive impacts on families and caregivers, during 40 Hz ISF therapy in a long-term care population. The high retention rate and consistently positive outcomes warrant larger-scale controlled studies. Caregiver assessment offers particular value in dementia care settings where self-report reliability is limited. If confirmed at scale, these findings support the potential of ISF therapy as a low-burden assistive technology with meaningful socioeconomic returns—reducing care demands, alleviating caregiver strain, and improving quality of life across the full care ecosystem, with implications for facilities, insurers, and public health systems managing the growing burden of dementia.

1 Introduction

Globally there are approximately 55 million people living with dementia, a number projected to triple by 2050 (Ismail et al., 2022). As the patients with dementia symptoms worsen and their ability to care and remain independent is limited, their family members, spouses, and acquaintances often become caretakers. Providing care for individuals living with dementia can have significant consequences on these caregivers, including decreased physical and mental health which includes increased rates of emotional distress and depression, and a higher risk of mortality than non-caregiving individuals (Pinquart and Sörensen, 2003; Sörensen et al., 2006; Liu et al., 2017). In the United States alone, more than 7 million people are living with Alzheimer's disease, with total annual care costs exceeding $360 billion—encompassing healthcare, long-term care services, and the estimated economic value of 18.4 billion hours of unpaid family caregiving (Alzheimer's Association, 2024). Within long-term care facilities, behavioral and psychological symptoms of dementia (BPSD)—including sleep disturbances, agitation, mood dysregulation, and cognitive disengagement—are the primary drivers of care intensity and staff workload (Schulz and Sherwood, 2008). These symptoms increase the need for one-to-one monitoring, disrupt communal activities, impair residents' ability to perform activities of daily living (ADL) independently, and contribute to self-reported social isolation, sleep deprivations, anxiety, and other well-documented crisis of caregiver burnout and staff turnover across the sector (Yu et al., 2015; Raggi et al., 2015; Liu et al., 2017).

Currently available interventions for managing BPSD range from pharmacological approaches (antipsychotics and sedatives that carry significant side effects, prolonged economic burden and limited long-term efficacy) to non-pharmacological strategies such as music therapy, reminiscence programs, and structured activity engagement. While non-pharmacological interventions show promise, they typically require substantial staff time for consistent delivery and remain underfunded. A critical gap exists for low-burden interventions that can improve resident engagement, functional status and quality of life without increasing staff demands.

40 Hz Invisible Spectral Flicker (ISF) therapy, delivered via the EVY LIGHT® device (Optoceutics, (2025)), represents a novel class of **assistive technology** designed to address this gap. Unlike conventional therapeutic interventions requiring active participation or staff facilitation, ISF therapy is passively administered: residents complete one hour of daily exposure in their own room during normal indoor activities, without staff involvement beyond initial device placement. Initial preclinical evidence has established that 40 Hz light-based gamma stimulation attenuates amyloid pathology and neuroinflammation in Alzheimer's disease models (Iaccarino et al., 2016; Gray et al., 1989; Herrmann, 2001), and combined with audiovisual stimulation has shown multi-region neuroprotective effects (Martorell et al., 2019; Ismail et al., 2022). Building on this preclinical evidence, a randomized, placebo-controlled pilot study confirmed the safety and potential clinical efficacy of ISF in mild-to-moderate Alzheimer's patients (Agger et al., 2023). Having established that the device is effective and well-tolerated in-home settings, the present study examines how ISF therapy performs in an institutional long-term care environment and, critically, whether its benefits extend beyond the individual resident to the caregivers, care staff, and families who constitute the broader care ecosystem.

This report presents a **real-world evidence (RWE) study**, reflecting both the practical realities of long-term care research and the growing recognition that RWE is essential for understanding how interventions can impact the neediest population. While Electronic Patient-Reported Outcome Measures (ePROMs) provide valuable insight into user-perceived therapeutic benefits, they present inherent limitations in populations with cognitive impairment, low tech-literacy, or advanced age. Self-report reliability may be compromised in individuals with dementia or mild cognitive impairment, who may have difficulty accurately recalling or assessing changes in their condition. Research on dementia-specific quality of life instruments has established that proxy-reported assessments by caregivers provide valid and reliable measures when self-report is limited (S. C. Smith et al., 2005). **Caregiver-assessed outcome measures** offer a complementary approach, providing third-party observation by trained professionals who interact with residents daily and can reliably detect subtle changes in behavior, mood, and functional status. As one caregiver noted during this study:

> *"Staff and myself have noticed a great change in resident. He is more pleasant and willing to perform simple tasks with the aides when asked. Definitely overall more engaging."* — Week 2 caregiver note

This report presents an analysis of caregiver-assessed outcomes from a pilot study conducted in partnership with **Masonicare**, Connecticut's largest not-for-profit senior care organization. At Masonicare at Mystic, CT, one of the organization's locations, 10 residents of the Memory Care unit received an EVY LIGHT® device. Between November 2025 and January 2026, each completed at least 1 hour of daily 40 Hz ISF therapy for 6 weeks. The primary objectives are to evaluate caregiver-observed changes during the initial six weeks of therapy, to assess the impact on resident quality of life, quantify changes in caregiver burden, and family wellbeing, and to characterize the potential socioeconomic implications of broad deployment in long-term care settings. The dataset comprises **280 individual assessments** collected from **10 residents**, across **5 domains** and spanning **6 weeks**. Different residents started the therapy at different times, but they were evaluated every week since they started the therapy.

The assessment protocol consists of a weekly five-item questionnaire completed by trained caregivers, evaluating observed changes from baseline in Sleep, Focus, Mood, Energy, and Memory. Caregivers report the resident's status as "Worse," "Stable," or "Improved." To facilitate quantitative analysis, these categorical responses are mapped to algebraic values (-1, 0, $+1$), allowing for the calculation of cumulative improvement scores and average improvement rates across the cohort.

A key strength of this pilot study is the exceptional retention rate:

9 of 10 participants (90%) completed all six weeks of assessment. This high completion rate contrasts favorably with typical voluntary digital health studies and provides near-consistent longitudinal data (n=10 for weeks 1–2; n=9 for weeks 3–6 due to one withdrawal for reasons unrelated to the study) without the confounding effects of substantial participant attrition (Eysenbach, 2005; Baumel et al., 2019). This analysis aims to characterize caregiver-observed therapeutic response patterns, identify domains with the strongest observed improvements, quantify the reduction in caregiver burden, and provide preliminary evidence supporting both the clinical and socioeconomic case for larger-scale controlled investigations.

2 Results

2.1 Data acquisition

Weekly caregiver assessments were collected from Masonicare at Mystic, a long-term care facility in the United States. Unlike self-reported ePROMs, these assessments were completed by trained caregivers who observed residents throughout their daily activities, providing an objective third-party perspective on resident status.

Caregivers were asked weekly, for the first 6 weeks after residents started the light therapy, to complete a 5-question assessment. This evaluates observed changes from baseline in 5 areas: **Sleep**, **Focus**, **Mood**, **Energy** and **Memory**. For each question caregivers select **Worse**, **Stable** or **Improved** based on their observations.

For the analysis, an “Improved” answer counts as “+1” and a “Worse” answer counts as “-1”, while “Stable” counts as “0”, enabling us to quantify categorical data and perform algebraic operations. There are two approaches when analyzing these responses: one where we treat each assessment as independent across residents, weeks and questions as either -1, 0 or $+1$, and another where we cumulatively add subsequent “Improved” counts or subtract the “Worse” counts, aggregated by resident, weeks or question types.

2.2 Data description

The data was gathered between November 2025 and January 2026. **10** residents from the memory care unit at Masonicare started light therapy during this period were enrolled in the pilot study and had caregiver assessments completed. Notably, **9 of 10 residents (90%)** had assessments completed for all 6 weeks, demonstrating exceptional retention compared to typical voluntary health studies. The one resident who dropped out did so for reasons unrelated to the study, but their first two weeks of data were still included in the analysis.

In total, **56 weekly assessments** were completed (10 residents × 6 weeks, minus assessments for the one resident who was withdrawn after week 2), yielding **280 individual question responses** across the five domains. Unless specified, the results include the first two weeks of data for the resident who withdrew. Therefore, n varies from 10 at weeks 1 and 2, to 9 at weeks 3 to 6.

2.2.1 Participants

Demographic and clinical characteristics of the 9 residents who completed the 6-week treatment are summarized in Fig.1. The sample was predominantly female (7 of 9) residents with complete demographic data whose primary diagnoses including Alzheimer’s disease, vascular dementia, mixed dementia, and other/unspecified dementia. Most residents were classified at Stage 5 (Moderately Severe) on the dementia staging scale, with the two male residents at Stage 6 (Severe), out of a maximum of 7, as per the Global Deterioration Scale, (Reisberg, Vadukapuram, and Kenowsky, 2022).

Fig.2 shows the distribution of time since diagnosis. The variable duration since diagnosis (approximately 0.5 to 2 years) reflects the heterogeneous disease trajectories typical of long-term care populations and provides context for interpreting both baseline cognitive status and potential responsiveness to intervention. The two residents at a more advanced stage of dementia, were also some of the first to receive the diagnosis.

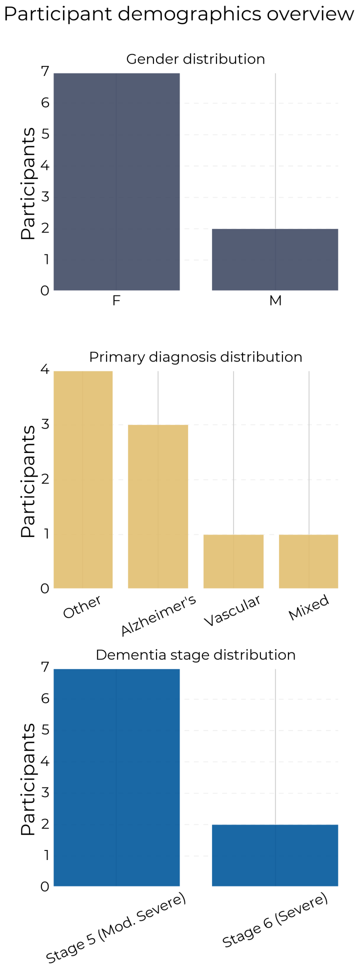

Figure 1: Participant demographics: distribution of biological gender, primary diagnosis, and dementia stage (n=9 with complete data).

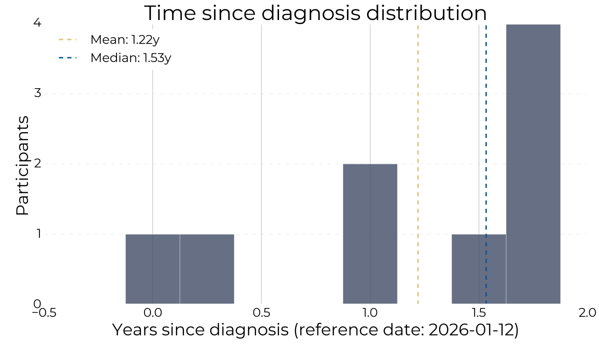

Figure 2: Distribution of years since dementia diagnosis, with mean and median times since diagnosis

2.2.2 *Assessments per week*

Looking at the distribution of caregiver responses per week, we produce a plot like Fig.3, normalized to the number of assessments of that week. A striking pattern emerges: while week 1 showed predominantly "Stable" assessments (82.0%) with few "Improved" (8.0%) and some "Worse" (10.0%), week 2 marks a clear transition with "Improved" assessments rising to 46.0%. From week 3 onwards, the "Improved" rate stabilizes between 60-67%, and remarkably, **zero "Worse" assessments** were recorded from week 3 through week 6. This consistent positive trajectory suggests a therapeutic response pattern where initial stabilization gives way to observable improvement.

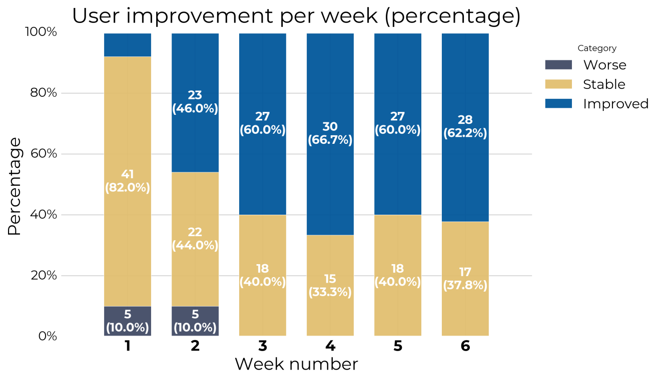

Figure 3: Percentage of each assessment option per week, across all 5 domains

2.2.3 *Assessments per question type*

Looking at the assessments per question type, across all weeks, we produce a plot like Fig.4. We define **Improvement** as the difference between the number of "Improved" minus "Worse" assessments, Eq.1. This could essentially be called the **net improvement**. These metrics can be aggregated for each resident, for each week, for each question, or a combination of these.

The **Average Improvement** is calculated as the number of "Improved" assessments minus the number of "Worse" assessments divided by the total number of assessments (see Eq.2).

$$\text{Improvement} = \#\text{Improved} - \#\text{Worse} \tag{1}$$

$$\text{Avg. imp. } (\%) = \frac{\#\text{Improved} - \#\text{Worse}}{\#\text{Improved} + \#\text{Stable} + \#\text{Worse}} \times 100\% \tag{2}$$

Alternatively, the bars can be stacked to produce the plot in Fig.5.

A key finding is that **all five domains show positive net improvement**. Mood and Sleep show the highest average improvement rates (+67.9% and +64.3% respectively), followed by Focus (+44.6%), Energy (+28.6%), and Memory (+25.0%). This contrasts with self-reported ePROM studies in which Memory often shows negative or neutral outcomes; here, caregiver observation detected positive Memory changes, albeit at a lower rate than other domains. These improvements across these domains suggest meaningful enhancements to residents' daily functioning and overall quality of life.

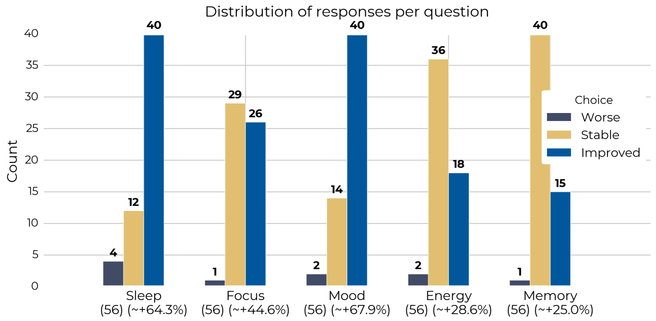

Figure 4: Count of each assessment option for each domain

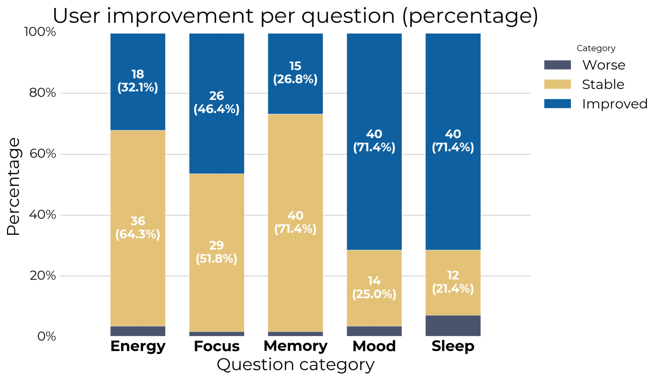

Figure 5: Percentage of each assessment option for each domain

2.3 Improvements per week and question

Going into further detail, we visualize the average improvement (Eq.1 and Eq.2) aggregated per week and per question type. The plots in Fig.6 and Fig.7 display essentially the same data, with the x axis and bars swapped. The first plot more clearly outlines the differences between question types, while the second more clearly outlines the trends through time.

In these plots, an average improvement of 0.50 in a week for a certain question means that if X% of caregivers assessed “Worse”, then X% + 50% assessed “Improved.”

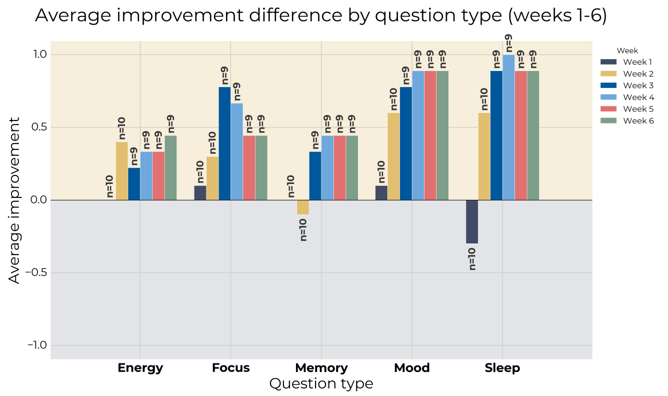

Figure 6: Average improvement in assessments per question and per week

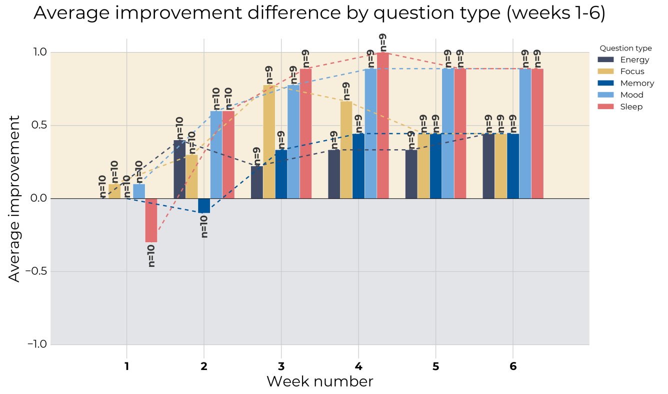

Figure 7: Average improvement in assessments per week and per question, with trace lines

Several trends are evident from these plots:

- **Sleep and Mood** consistently show the highest improvement rates across most weeks, with Mood reaching 1.0 (100% positive improvement) in week 4.
- **Week 1** shows near-zero or slightly negative improvement across all domains, consistent with an initial stabilization phase before observable therapeutic effects emerge.
- **Memory** shows a delayed onset pattern, with improvement becoming consistently positive from week 3 onwards.
- From the trace lines in Fig.7, all domains trend upward from week 1 to week 2, with most maintaining elevated improvement rates through week 6.

2.4 Improvement progression plots

This section analyzes the improvement metric longitudinally to track changes week over week. This study benefits from 90% retention from week 3 onwards, and the results from the resident who eventually withdrew were still included for the first 2 weeks.

Improvement was evaluated using two methods: **discrete** (weekly scores of -1, 0, or +1) and **cumulative** (a running total of scores over time up until week X) for each user. While the results with the discrete method range from -1 to +1, the cumulative results up to week X can range from -X to +X.

In these improvement progression plots, the thickness of the lines is proportional to the number of assessments for that week. The 95% confidence interval for the mean improvement is also included, obtained from the standard error of the mean $CI = \text{mean} \pm 1.96 \cdot SD/\sqrt{n}$ (see Methods sec.4.3.2).

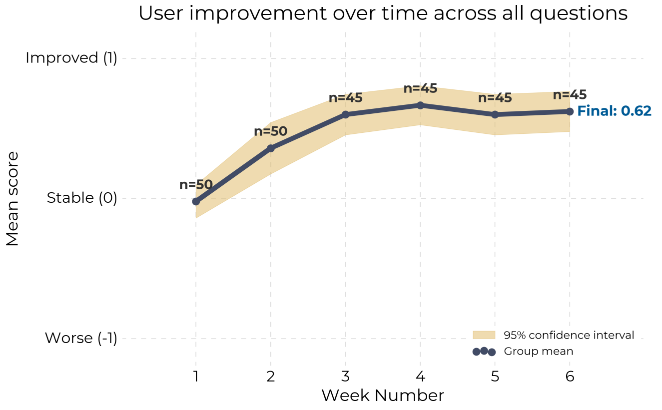

Figure 8: Caregiver-assessed average improvement progression plot for all questions, with confidence interval of the mean. Scale from -1 to +1.

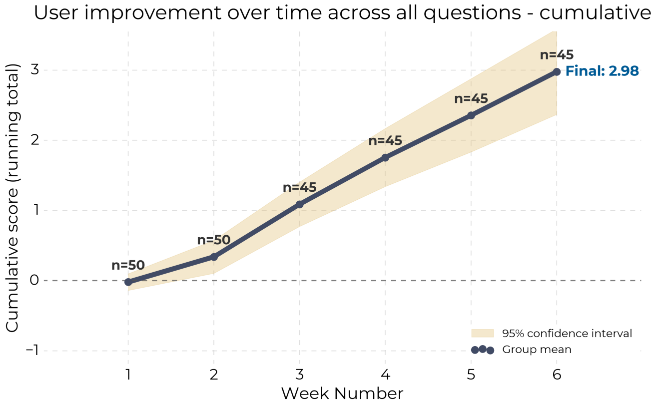

Figure 9: Cumulative caregiver-assessed improvement progression plot for all questions, with confidence interval of the mean. Scale from -6 to +6.

From these two plots we observe that, even though week-by-week improvements start near zero in week 1 (Fig.8), they rise sharply in week 2 and maintain positive values throughout. The cumulative effect is particularly striking (Fig.9): the average cumulative improvement rises steadily from near zero to **+2.98 by week 6**, with the 95% confidence interval fully above zero from week 3 onwards. This indicates, with high confidence, that the average resident experienced net positive improvement over the 6-week period.

2.4.1 *Question by question*

Over this subsection we examine each question type individually. The weekly improvement progression plots are included in the Appendix (Supplementary figures A). Here we focus on the **cumulative** improvements for each question.

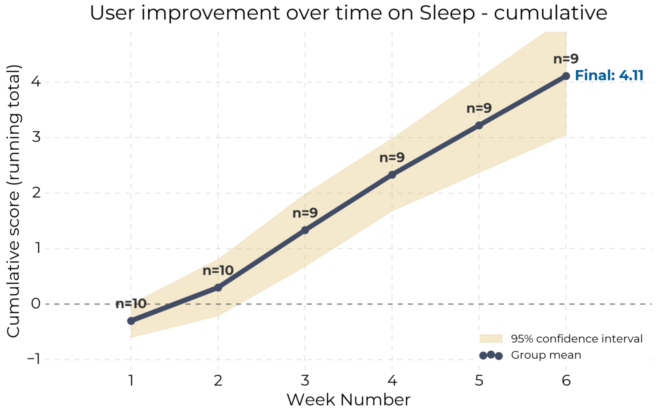

Figure 10: Cumulative caregiver-assessed improvement progression on Sleep

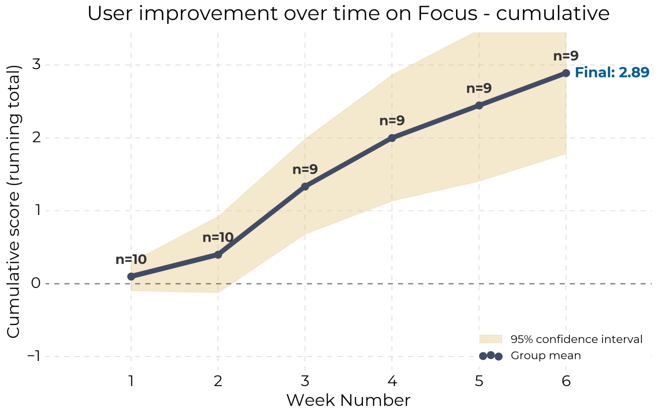

Figure 11: Cumulative caregiver-assessed improvement progression on Focus

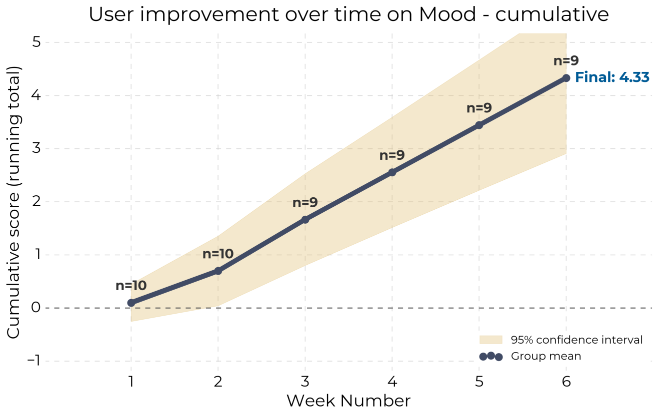

Figure 12: Cumulative caregiver-assessed improvement progression on Mood

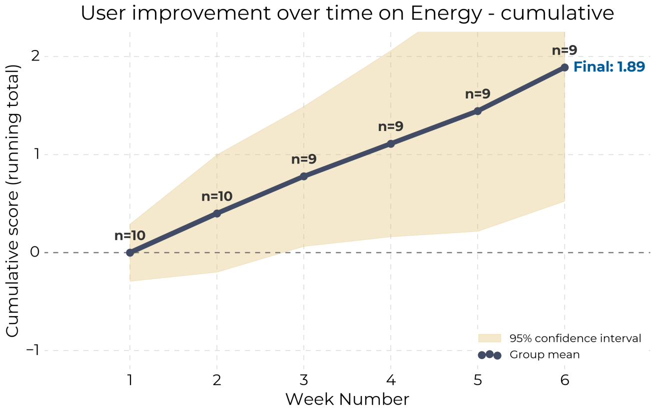

Figure 13: Cumulative caregiver-assessed improvement progression on Energy

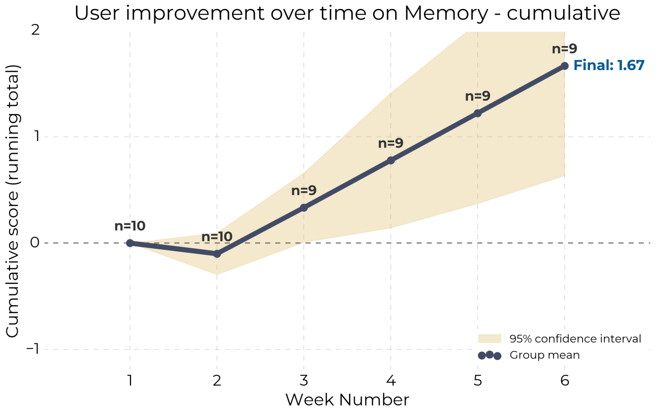

Figure 14: Cumulative caregiver-assessed improvement progression on Memory

Sleep (Fig.10) demonstrates strong and consistent week-on-week improvement, with the cumulative score rising steadily throughout the 6-week period. Caregivers noted tangible changes in residents' sleep patterns:

> *"Private aide reports that the resident sleep has improved greatly. When she first started working with the resident she was up every few hours."* — Week 5 caregiver note

These sleep improvements directly translate to reduced nighttime disruptions and decreased demands on overnight caregiving staff, alleviating a significant source of caregiver burden in long-term care settings.

Focus (Fig.11) and **Mood** (Fig.12) both show robust positive trajectories, with Mood exhibiting particularly strong improvement. Both domains have confidence intervals that remain above or near zero throughout most of the study period. One caregiver captured the behavioral changes associated with mood improvement:

> *"Over the past few weeks I have noticed that resident has been coloring in room, she hasn't done in many months. This was once her favorite pass time that*

> *gave her joy. It's great to see her revisiting her past interest!"* — Week 5 caregiver note

Energy (Fig.13) shows positive cumulative improvement, though with a slightly more variable trajectory than Sleep or Mood.

Memory (Fig.14) displays the characteristic delayed-onset pattern: the cumulative score dips slightly in week 2 before recovering and trending positive from week 3 onwards. This pattern is consistent with the hypothesis that cognitive domains may require longer exposure periods before observable improvements manifest. When improvements did emerge, caregivers reported meaningful changes:

> *"Resident has some moments of clarity, and is able to engage in meaningful conversation."* — Week 3 caregiver note

2.5 Absolute improvement analysis

While the progression plots above show *average* improvement per resident, the **absolute improvement** metric provides a different perspective by showing the raw count difference between "Improved" and "Worse" assessments each week. This is equivalent to weighting each assessment equally regardless of which resident it came from.

The absolute improvement is defined as:

$$\text{Absolute Improvement}_{\text{week}} = \#\text{Improved}_{\text{week}} - \#\text{Worse}_{\text{week}} \tag{3}$$

This metric is particularly valuable in this pilot study because it clearly visualizes the total therapeutic benefit observed across all residents.

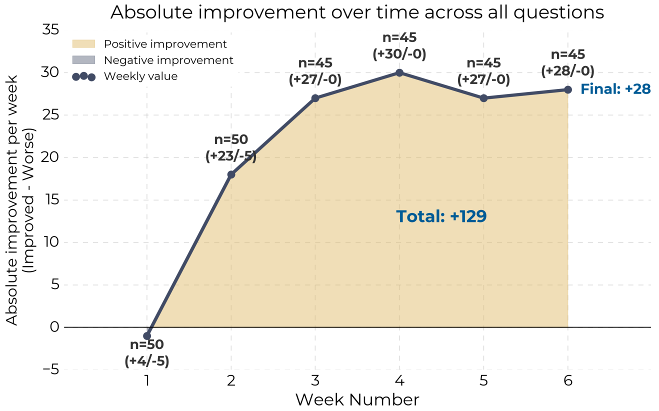

Figure 15: Absolute improvement (Improved minus Worse) per week across all questions. The shaded area represents the net accumulated benefit over time.

Fig.15 shows the week-by-week absolute improvement. Several key observations emerge:

- **Week 1** shows a slight negative absolute improvement (-1), with 4 "Improved" and 5 "Worse" assessments out of 50 total.
- **Week 2** marks a dramatic shift to $+18$ absolute improvement (23 "Improved", 5 "Worse").
- **Weeks 3–6** show consistently high positive values ($+27$ to $+30$), and crucially, **zero "Worse" assessments** were recorded during this entire period.
- The **Total absolute improvement across 6 weeks is +129**, representing a substantial net positive therapeutic observation.

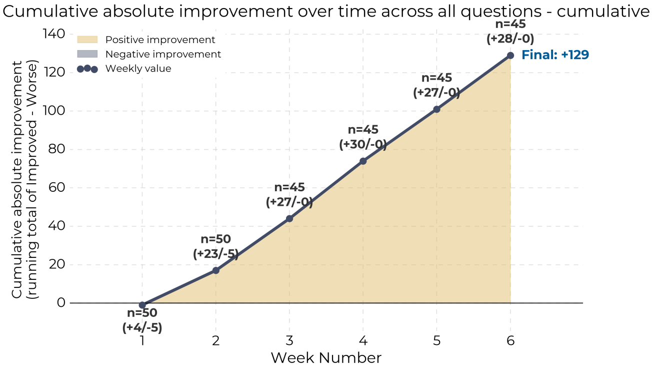

Figure 16: Cumulative absolute improvement (running total) across all questions

Fig.16 shows this same data as a running total, illustrating the steady accumulation of positive outcomes over the 6 weeks.

The absence of any "Worse" assessments from week 3 onwards is particularly noteworthy. While the small sample size ($n = 10$) limits statistical inference, this pattern suggests that after an initial stabilization period, residents either maintain their baseline status or show improvement—none were observed to decline.

2.5.1 Absolute improvement by domain

The absolute improvement area plots for individual domains are included in the Appendix (Supplementary figures C).

Of particular interest is the **Memory** domain, which shows a delayed onset pattern: neutral in week 1, slightly negative in week 2 (-1), then consistently positive from week 3 onwards with a total of $+14$ absolute improvement. This delayed response in Memory, compared to more immediate improvements in Mood and Sleep, is consistent with the hypothesis that cognitive domains may require extended exposure periods before therapeutic effects become observable.

2.6 Cumulative improvement progression with trajectories

The cumulative improvement progression plots can be visualized with trajectory plots, which show the improvement progression for each user. These plots are included in the Appendix (Supplementary figures B). The thickness of the blue lines is proportional to the number of users who have that specific improvement progression. These plots are equivalent to the cumulative improvement progression plots, but with the individual trajectories for each user overlaid.

One key take away is that most trajectories are of constant improvement or stabilization after week 3 or 4. At the same time, no user ever gets a cumulative improvement below minus 1, meaning that even when residents get worse, they don't continue getting worse and eventually stabilize.

2.7 Statistical significance testing

To formally assess whether the observed improvements represent statistically significant changes from baseline, we performed paired comparisons between week 1 (baseline) and subsequent weeks using the Wilcoxon signed-rank test. Effect sizes were quantified using Cohen's d and Cliff's delta.

For each comparison, we matched residents who had assessments in both the baseline week and the target week, enabling paired statistical testing. Despite the small sample size ($n = 9$ for weeks 3–6), several comparisons reached statistical significance with large effect sizes.

2.7.1 *Global improvement*

The global (all domains combined) improvement showed statistically significant differences from baseline starting at **week 3** (Wilcoxon $p = 0.004$, Cohen's $d = 2.40$). This significance was maintained through week 6 ($p = 0.004$, $d = 2.16$). The effect sizes are classified as "large" (Cohen's $d > 0.8$), indicating substantial practical significance.

2.7.2 *Domain-specific significance*

Table 1 summarizes the key statistical findings by domain:

Domain	1st Sig. Week	Cohen's d	Cliff's δ	p-value	Effect Size
Global	Week 3	2.40	0.86	0.004	Large
Sleep	Week 2	1.50	0.69	0.021	Large
Mood	Week 3	1.27	0.58	0.014	Large
Focus	Week 3	1.71	0.67	0.014	Large
Memory	Week 4	1.19	0.44	0.046	Large
Energy	(W. 6)	(0.87)	(0.40)	n.s.	Large

Table 1: Statistical significance of improvement from baseline by domain. "First Sig. Week" indicates the earliest week at which $p < 0.05$ was achieved. Cohen's d, Cliff's δ and p-value are reported for the first significant week. All significant results show large effect sizes (given by Cohen's $d > 0.8$). (n.s. = not significant)

Key observations:

- **Sleep** reaches significance earliest (week 2), consistent with caregiver observations that sleep improvements were among the first noticeable changes.
- **Mood** and **Focus** showed significant improvements by week 3, with Mood maintaining significance through week 6.
- **Memory** shows the delayed onset pattern noted earlier, reaching significance only at week 4—consistent with the hypothesis that cognitive domains require longer exposure.
- **Energy** does not reach statistical significance at any timepoint, though it shows positive trends. This is consistent with Energy having the lowest improvement rate (28.6%). Despite achieving a higher average improvement score than Memory, Energy does not reach significance mostly because of the big variance in Energy improvement among users (see Fig.13).

By week 6, the comparison with baseline shows:

- Global: $d = 2.16$, $p = 0.004$ (highly significant)
- Sleep: $d = 2.88$, $p = 0.004$ (highly significant)
- Mood: $d = 1.60$, $p = 0.008$ (highly significant)
- Memory: $d = 1.19$, $p = 0.046$ (significant)

These results demonstrate that despite the small sample size, the observed improvements are not merely descriptive trends but represent statistically robust changes with large effect sizes.

2.8 Staff burden reduction

To quantify the staff burden reduction of the intervention, the caregiver, which was the same for each resident, completed a small survey. This tool is a simple assessment with 5-point Likert scales that evaluates the burden reduction of the caregiver with respect to each resident, across 4 domains.

Each domain was rated at baseline (week 1) and at final (week 6) on a 5-point Likert scale. Anchors were: 1 = Strongly Negative (severe burden, high stress, constant strain); 2 = Somewhat Negative (notable burden, elevated stress); 3 = Neutral (no meaningful impact on caregiver experience); 4 = Somewhat Positive (reduced burden, lower stress); 5 = Strongly Positive (significant relief, marked improvement in caregiver experience). For analysis and visualization, scores were re-centered to a -2 to $+2$ scale with 3 (Neutral) mapped to 0, so that zero represents no change and deviations from zero correspond to improvement or worsening.

The four domains evaluated reflected the impact of taking care of a particular resident in the personal and professional life of the caregiver. These were:

- **Emotional Exhaustion** — Level of stress, anxiety, or emotional fatigue experienced when caring for the resident.
- **Sense of Control** — How prepared, capable, and in-control the caregiver felt during shifts with the resident.
- **Team Resource Strain** — How often backup staff, escalation, or additional support was needed for the resident.
- **Overall Job Impact** — The effect caring for the resident had on overall morale and willingness to continue in the role.

Fig.17 shows the composite burden score (average across the four domains) for each resident from baseline to final, along with the group mean ± 95% confidence interval. All but one trajectory moves upward, indicating reduced burden or improved caregiver experience. Fig.18 presents the four domains individually; each shows positive mean change from baseline to final, with the zero line (no change) clearly below the final mean.

One particular participant showed no change in caretaker burden on any of the 4 domains. On the other hand, all other showed improvements ranging from +1 to +3 points on the composite score. The 2 participants with the more advanced dementia stage, who were associated with a lower staff burden at baseline, showed the largest improvements, of +2 and +3 points on the composite score.

Table 2 summarizes the mean baseline, final, and change for each domain. Emotional Exhaustion showed the largest mean improvement (+1.67), followed by Team Resource Strain and Overall Job Impact (+1.56 each), and Sense of Control (+1.33). All domains show a statistically significant improvement from baseline to final, with paired 2-sided t-tests yielding p-values < 0.05 for each domain and for the composite score, for $n = 9$.

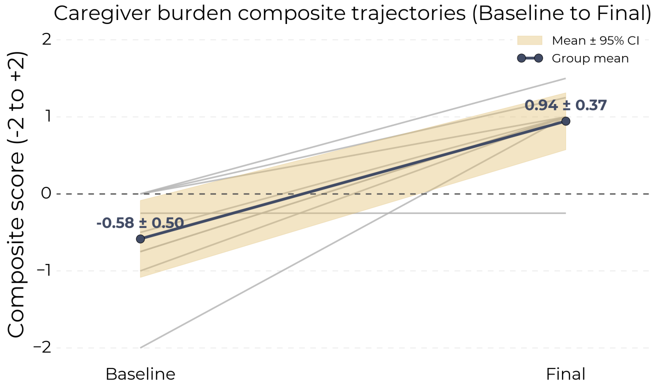

Figure 17: Average caregiver burden trajectories from baseline to final. Individual participant lines (gray), group mean $\pm$ 95% confidence interval (blue/gold). Scores on -2 to $+2$ scale (3 = no change).

Metric	Avg Baseline	Avg Final	Avg Change	p-val
Composite average	-0.58	0.94	+1.52	0.012
Emotional Exhaustion	-0.78	0.89	+1.67	<0.01
Sense of Control	-0.44	0.89	+1.33	<0.01
Team Resource Strain	-0.56	1.00	+1.56	0.010
Overall Job Impact	-0.56	1.00	+1.56	0.010

Table 2: Staff burden four-domain assessment: mean baseline (week 1), final (week 6), and change. Scale from -2 to +2; positive change reflects improvement in caregiver experience. P-values are for the Wilcoxon signed-rank test between baseline and final,for $n = 9$.

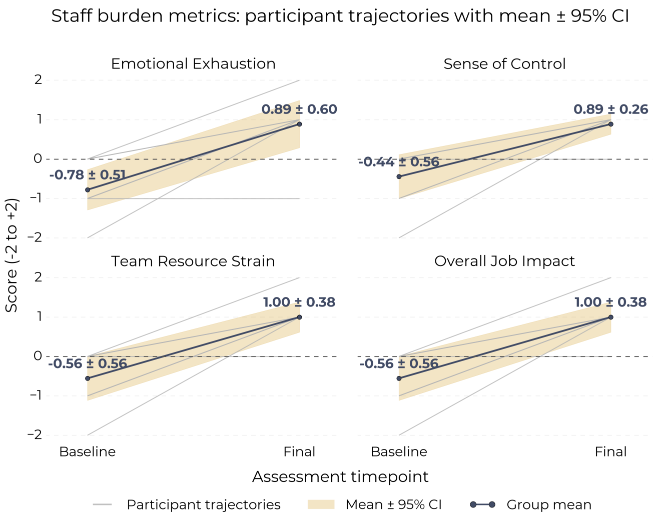

Figure 18: Staff burden metrics by domain: individual trajectories (gray), and group mean $\pm$ 95% confidence interval (blue/gold) from baseline to final. Scores on -2 to $+2$ scale (3 = no change).

3 Discussion

This pilot study analyzed caregiver-assessed outcome measures collected from 10 residents at Masonicare, a US-based long-term care facility, over a 6-week period of 40 Hz Invisible Spectral Flicker therapy. The findings provide preliminary evidence supporting multi-domain improvements in residents as observed by trained caregivers and improved caretaker burden via caretaker self-assessment measures.

3.1 Key findings

- **All five assessed domains showed positive net improvement**: Mood (+67.9%), Sleep (+64.3%), Focus (+44.6%), Energy (+28.6%), and Memory (+25.0%).
- **Exceptional retention**: 9 of 10 residents (90%) completed all 6 weeks of assessment, providing consistent longitudinal data without the confounding effects of substantial participant attrition.
- **Sustained positive trajectory**: From week 3 onwards, caregivers reported zero "Worse" assessments across all domains. This indicates that after an initial stabilization period, residents either maintained their baseline status or showed continuous observable improvement.
- **Strong cumulative benefit**: The total absolute improvement across 6 weeks was +129 (sum of "Improved" minus "Worse" assessments). The average cumulative improvement score reached +2.98 per resident by week 6, with the 95% confidence interval fully above zero from week 3 onwards.
- **Statistical significance with large effect sizes**: Despite the small sample size ($n = 9$), global improvement reached statistical significance by week 3 (Wilcoxon $p = 0.004$,

Cohen's $d = 2.40$). Sleep showed the earliest significance (week 2, $p = 0.021$), while Memory exhibited delayed significance (week 4, $p = 0.046$). All significant comparisons showed large effect sizes (Cohen's $d > 1.0$).

- **Delayed onset in Memory**: While Mood and Sleep showed rapid improvement from week 2, Memory exhibited a delayed response pattern with consistent positive observations emerging from week 3. This is consistent with the hypothesis that cognitive domains may require extended exposure periods before therapeutic effects become observable.
- **Staff burden reduction**: Post-intervention caregiver burden assessment (Section 2.8) showed an average burden reduction score of +1.52 (on a scale of -2 to $+2$). Eight of nine residents (89%) showed substantial reduction, with a change from baseline $\geq +1$, across the four domains—Emotional Exhaustion, Sense of Control, Team Resource Strain, and Overall Job Impact. All residents showed stabilization or improvement; none showed worsening burden. The 2 participants with the more advanced dementia stage, who were associated with a lower staff burden at baseline, showed the largest improvements.

Beyond the quantitative measures, caregiver notes captured the human impact of these improvements. Family members observed meaningful changes:

> *"Son states he is in a better mood and less anxious."*
> — Week 5 caregiver note

3.2 Advantages of caregiver assessment

The caregiver-assessed methodology employed in this study offers several advantages over self-reported ePROMs, particularly in populations with cognitive impairment:

- **Reduced self-report bias**: Individuals with dementia or mild cognitive impairment may have difficulty accurately recalling or assessing changes in their own condition. Caregiver observation provides an objective third-party perspective.
- **Consistency of assessment**: Trained caregivers who interact with residents daily can detect subtle changes in behavior, mood, sleep patterns, and functional status that residents themselves may not perceive or report.
- **Reliability in longitudinal studies**: The high retention rate (90%) suggests that caregiver-completed assessments may be more sustainable than patient self-report in long-term care settings.

Notably, the Memory domain showed positive improvement in caregiver assessments (+25.0%), whereas self-reported ePROM studies have often shown negative or neutral Memory outcomes. This discrepancy may reflect the limitations of self-perception in cognitive domains, or the ability of caregivers to detect functional improvements that residents themselves do not recognize.

3.3 Socioeconomic impact

Beyond the quantitative clinical improvements, this pilot study revealed meaningful impacts on residents' quality of life, family experiences, and the broader care delivery environment. Multiple family members independently observed and reported improvements in their loved ones. Daughters noted that residents were "in a better mood" (Week 5 observations from three separate families), while one son specifically stated his family member was "in a better mood and less anxious" (Week 5). Family involvement extended beyond observation: in one case, a daughter actively reassured her mother about the therapy's benefits (Week 2), demonstrating family engagement in supporting the intervention.

The improvements observed by caregivers translated into tangible quality of life enhancements for residents. Several residents returned to meaningful activities they had abandoned: one resident resumed coloring "after many months"—an activity that "was once her favorite pass time that gave her joy" (Week 5); another began "wearing lipstick daily, something she hasn't done in months" (Week 6); and a third returned to reading books, "once a past interest that has been revisited" (Week 5). Social engagement also increased: residents attended meals in common areas without prompting, participated in activities they had previously shown "little or no interest in," and one resident attended a Christmas movie "that we haven't seen her do in over a year" (Week 2). These observations suggest residents were not merely showing clinical improvement on assessment scales, but were reclaiming aspects of their identity and experiencing more fulfilling daily lives.

The reduction in caregiver burden documented in Section 2.8 has broader implications for care quality and staff wellbeing. When residents sleep better, caregivers face fewer nighttime disruptions; when residents are more cooperative and pleasant, staff spend less time on behavioral management and more time on meaningful interactions; when residents require less intensive monitoring, staff can allocate their attention more effectively across their responsibilities. This decreased burden is particularly significant given the well-documented crisis of caregiver burnout in long-term care settings (Schulz and Sherwood, 2008). Reducing staff burden may contribute to improved caregiver health and longevity in the profession, potentially addressing the burnout and retention challenges that compromise care quality.

These resident-level and caregiver-level improvements carry direct implications for the economics of long-term dementia care—a sector under severe and growing financial strain. In the United States, total annual costs of caring for individuals with Alzheimer's and related dementias exceed \$360 billion, with long-term care services representing the largest single component (Alzheimer's Association, 2024). Approximately two-thirds of nursing home residents with dementia are covered by Medicaid, meaning that any reduction in care intensity translates directly into public fiscal savings. Behavioral and psychological symptoms of dementia—the very domains showing the greatest improvement in this study (Mood, Sleep)—are among the costliest to manage as they drive emergency department visits, require one-to-one staffing, and accelerate transitions to higher levels of care.

If the improvements observed in this pilot study scale to larger populations, the return on investment (ROI) for 40 Hz ISF therapy as assistive technology becomes compelling across multiple stakeholders:

- **Care staff**: A mean improvement of +1.52 points, in a 5 point scale, on the composite burden score (Section 2.8) across 4 domains, representing a substantial reduction in caregiver burden, directly reduces the risk of staff burnout and turnover. Replacing a single certified nursing aide costs an estimated $3,000–$5,000, and nursing staff replacement costs are substantially higher; reducing attrition represents a meaningful and direct facility cost saving (Alzheimer's Association, 2024).
- **Care facilities**: Reduced behavioral events and nighttime disruptions translate to measurable savings in staff hours and overtime. Improved resident engagement and quality-of-life outcomes may also differentiate facilities in the long-term care market.
- **Families**: Reduced anxiety about loved ones' wellbeing, fewer family interventions prompted by behavioral crises, and the profound psychological relief of seeing a loved one re-engage with daily life represent significant but often unmeasured forms of value.
- **Insurers and Medicare**: Fewer BPSD-driven hospitalizations and emergency department visits, and potential delay in transitions to higher-acuity care, represent meaningful cost avoidance for payers in the healthcare system.
- **State Medicaid programs and the public health system**: As the primary funder of long-term care for low-income individuals, states bear the largest fiscal burden of advanced dementia; any intervention that slows functional decline, reduces care intensity, or delays institutionalization at a higher level of care has direct budgetary implications.

Critically, the assistive technology model of ISF which delivers a passive, low burden, requiring no dedicated staff time during administration, means that these potential savings arise without adding to the workload of a strained care workforce, thereby reducing staff burden. The economic case is not simply that the therapy reduces costs; it is that it does so while simultaneously improving resident quality of life, alleviating professional caregiver burden, reassuring families, and generating evidence that can justify continued investment. While this pilot study cannot make definitive claims about long-term outcomes or aggregate cost savings, the observed pattern of resident improvements, reduced caregiver strain, and family-reported quality-of-life gains suggests that passively delivered assistive technologies targeting neurological engagement may yield benefits that cascade throughout the care ecosystem, creating the conditions for a virtuous cycle between resident wellbeing, staff health, and care quality.

3.4 Limitations

Several limitations should be acknowledged:

- **Small sample size**: With only 10 residents, statistical power is limited. The findings should be interpreted as preliminary and hypothesis-generating rather than definitive.
- **Absence of control group**: This observational pilot study did not include a placebo or no-treatment control arm. Observed improvements cannot be definitively attributed to the intervention versus placebo effects, regression to the mean, or natural fluctuation.
- **Potential observer bias**: Caregivers were not blinded to the intervention, which may introduce positive expectation bias in their assessments.
- **Single facility**: All participants were from one care facility, limiting generalizability to other populations or settings.
- Possible **seasonal confound**: The study was conducted during the winter holiday season, which may have introduced seasonal social-stimulation effects as a confound for early-week outcomes.

3.5 Future prospects

This preliminary study provided powerful insights and warrants larger-scale controlled studies. Future work to build upon this pilot would be to first increase the scale — from n=10, to n=40 residents at Masonicare at Mystic, possibly up to n=5000 in the total Masonicare network — and then the scope — with direct-to-consumer Activities of Daily Living (ADL) impact assessment.

This study showed remarkably positive results and response from caregivers and families, which is motivating for the use of EVY LIGHT® as assistive technology for long-term care residents. However, an increase in scale would allow for more powerful statistical analysis and generalizability regarding any of the metrics and claims. For instance, in the present study, the two participants with the more advanced dementia stage, who were associated with a lower staff burden at baseline, showed the largest improvements, of +2 and +3 points on the average score. However, with the small sample size, it is hard to correlate whether using EVY LIGHT® is the cause of the largest improvement in those residents, therefore warranting larger-scale studies to determine if this could be a stronger pattern.

3.6 Conclusions

This pilot study provides preliminary caregiver-assessed evidence supporting multi-domain improvements in resident quality of life with positive impacts on families and caregivers during 40 Hz Invisible Spectral Flicker therapy in a long-term care population. The key strengths of this study include:

- High retention rate (90%) providing consistent longitudinal data
- All five assessed domains showing positive net improvement
- Zero "Worse" assessments from week 3 onwards
- Statistically significant improvements with very large effect sizes — Cohen's $d > 2.0$ for global average and Sleep from week 3 onwards — despite small sample size.
- Meaningful reduction in caregiver burden: The 4-domain assessment showed consistent, statistically significant, positive change from baseline to final, with mean improvements of +1.52 points on the -2 to +2 scale. This reinforces that resident improvements translated into reduced caregiver strain across multiple dimensions. Multiple family members independently observed and reported improvements in their loved ones' mood and engagement.

Perhaps most compelling are the qualitative observations from caregivers who witnessed residents re-engaging with activities they had abandoned:

> *"Resident has been wearing lipstick daily, something she hasn't done in months."* — Week 6 caregiver note

> *"Resident has been coming to meals in the common area for breakfast without being prompted. She did attend a Christmas movie that we haven't seen her do in over a year."* — Week 2 caregiver note

Caregiver assessment may offer particular value in dementia care settings where self-report reliability is limited. The observed reduction in caregiver burden—reflecting less time spent on direct care, improved resident cooperation, and fewer sleep-related disruptions—suggests that resident improvements may translate into tangible benefits for care delivery and staff wellbeing. The promising results from this small pilot study—combining statistically significant quantitative improvements with meaningful qualitative observations and burden reduction—warrant larger-scale controlled investigations with randomized designs, blinded assessment, and longer follow-up periods to establish both clinical efficacy and real-world benefits for residents, families, and care staff.

4 Methods

This pilot was conducted as a structured quality improvement initiative within Masonicare at Mystic's existing care operations, generating real-world evidence on EVY LIGHT®'s patient-related outcomes and caregiver burden.

4.1 Data collection

Caregiver-assessed outcome data were collected in partnership with Masonicare at Mystic, a long-term care facility in the United States. This pilot study employed trained caregivers to complete weekly assessments of residents receiving 40 Hz Invisible Spectral Flicker therapy, providing caregiver-assessed resident outcome measures.

Caregivers completed a standardized 5-question assessment weekly for each resident, evaluating observed changes from baseline in Sleep, Focus, Mood, Energy, and Memory. For each domain, caregivers selected one of three categorical responses: "Worse," "Stable," or "Improved."

4.1.1 *Participant characteristics*

The pilot study enrolled 10 residents from Masonicare who began 40 Hz Invisible Spectral Flicker therapy. The data collection period spanned November 2025 through January 2026, with assessments completed over 6 consecutive weeks per resident.

Of the 10 enrolled residents, 9 (90%) had complete assessments for all 6 weeks. One resident was withdrawn after week 2 due to reasons unrelated to the study. This exceptionally high retention rate provides consistent longitudinal data without substantial attrition bias.

4.2 Data processing

4.2.1 *Score conversion*

Categorical caregiver responses were converted to numerical scores for quantitative analysis:

- "Improved" $\rightarrow +1$
- "Stable" $\rightarrow 0$
- "Worse" $\rightarrow -1$

This algebraic encoding enables calculation of cumulative improvement scores, average improvement rates, and absolute improvement metrics as described in Section 2.2.

On caregiver burden analysis, the 5 item Likert scales used were converted to a 5-point scale from -2 to +2. Positive values indicate reduction in burden and zero indicates no change in burden from baseline.

4.3 Analysis methods

4.3.1 *Descriptive statistics*

Improvement rates were calculated as described in equations 1 and 2. Both weekly (discrete) and cumulative improvement metrics were computed. Improvement and average improvement were aggregated by user, by week, by question, or a combination of these, for the different analyses.

4.3.2 Confidence intervals

95% confidence intervals for mean improvement scores and for the caregiver burden trajectories were calculated using the standard error of the mean with 95% confidence:

$$CI = \bar{x} \pm 1.96 \cdot \frac{s}{\sqrt{n}} \tag{4}$$

where $\bar{x}$ is the sample mean, s is the sample standard deviation, and n is the sample size. This assumed the samples followed normal distributions, and a deviation of 1.96 standard deviations from the mean includes 95% of the data in the population.

4.3.3 Statistical significance testing

To assess whether observed improvements represent statistically significant changes from baseline, paired comparisons were performed between week 1 (baseline) and subsequent weeks (2–6) using:

- **Wilcoxon signed-rank test**, Wilcoxon (1945): A non-parametric test for paired samples, appropriate for ordinal data with small sample sizes. Tests whether the median difference between paired observations differs significantly from zero.
- **Cohen's** d, Cohen (1988): A standardized measure of effect size calculated as the difference between means divided by the pooled standard deviation. Interpretation: $|d| < 0.2$ (negligible), $0.2 \leq |d| < 0.5$ (small), $0.5 \leq |d| < 0.8$ (medium), $|d| \geq 0.8$ (large).
- **Cliff's delta**, Cliff (1993): A non-parametric effect size measure for ordinal data, representing the probability that a randomly selected observation from one group exceeds a randomly selected observation from the other group, minus the reverse probability. Because the population distribution does not follow a normal distribution, the Cliff's delta is a more accurate measure of effect size. Magnitude thresholds follow Romano et al. (2006): $|\delta| < 0.147$ (negligible), < 0.33 (small), < 0.474 (medium), otherwise large.

For each comparison, only residents with assessments in both the baseline and target weeks were included, enabling valid paired testing.

4.4 Software and tools

Data analysis was performed using Python (`Python v3.9.13`) with the following key packages:

- `pandas v1.5.3` for data table manipulation and aggregation
- `numpy v1.24.4` for numerical computations
- `matplotlib v3.7.5` and `seaborn v0.13.2` for visualization
- `scipy v1.13.1` for statistical calculations

Analysis scripts and the data transformation pipeline are available in the project repository.

References

[1] Rawan Ismail et al. "An update on the use of gamma (multi)sensory stimulation for Alzheimer's disease treatment". In: *Frontiers in Aging Neuroscience* 14 (2022), p. 1095081. DOI: 10.3389/fnagi.2022.1095081.

[2] Martin Pinquart and Silvia Sörensen. "Differences between caregivers and noncaregivers in psychological health and physical health: A meta-analysis". In: *Psychology and Aging* 18.2 (2003), pp. 250–267. DOI: 10.1037/0882-7974.18.2.250.

[3] Silvia Sörensen et al. "Dementia care: mental health effects, intervention strategies, and clinical implications". In: *The Lancet Neurology* 5.11 (2006), pp. 961–973. DOI: 10.1016/s1474-4422(06)70599-3.

[4] Shuai Liu et al. "Caregiver burden and prevalence of depression, anxiety and sleep disturbances in Alzheimer's disease caregivers in China". In: *Journal of Clinical Nursing* 26.9-10 (2017), pp. 1291–1300. DOI: 10.1111/jocn.13601.

[5] Alzheimer's Association. "2024 Alzheimer's disease facts and figures". In: *Alzheimer's & Dementia* 20.5 (2024), pp. 3708–3821. DOI: 10.1002/alz.13809.

[6] Richard Schulz and Paula R Sherwood. "Physical and mental health effects of family caregiving". In: *The American Journal of Nursing* 108.9 Suppl (2008), pp. 23–27. DOI: 10.1097/01.NAJ.0000336406.45248.4c.

[7] Hongmei Yu et al. "Measuring the caregiver burden of caring for community-residing people with Alzheimer's disease". In: *PLOS ONE* 10.7 (2015), e0132168. DOI: 10.1371/journal.pone.0132168.

[8] Alberto Raggi et al. "The burden of distress and related coping processes in family caregivers of patients with Alzheimer's disease living in the community". In: *Journal of the Neurological Sciences* 358.1-2 (2015), pp. 77–81. DOI: 10.1016/j.jns.2015.08.024.

[9] Dec. 2025. URL: https://optoceutics.com/.

[10] Hunter F Iaccarino et al. "Gamma frequency entrainment attenuates amyloid load and modifies microglia". In: *Nature* 540.7632 (2016), pp. 230–235.

[11] Charles M Gray et al. "Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties". In: *Nature* 338.6213 (1989), pp. 334–337. DOI: 10.1038/338334a0.

[12] Christoph S Herrmann. "Human EEG responses to 1–100 Hz flicker: resonance phenomena in visual cortex and their potential correlation to cognitive phenomena". In: *Experimental Brain Research* 137.3–4 (2001), pp. 346–353. DOI: 10.1007/s002210100682.

[13] Anthony J Martorell et al. "Multi-sensory gamma stimulation ameliorates Alzheimer's-associated pathology and improves cognition". In: *Cell* 177.2 (2019), 256–271.e22. DOI: `10.1016/j.cell.2019.02.014`.

[14] Mikkel Pejstrup Agger et al. "Safety, feasibility, and potential clinical efficacy of 40 Hz invisible spectral flicker versus placebo in patients with mild-to-moderate Alzheimer's disease: a randomized, placebo-controlled, double-blinded, pilot study". In: *Journal of Alzheimer's Disease* 92.2 (2023), pp. 653–665.

[15] Sarah C Smith et al. "DEMQOL: A new patient-reported outcome measure for dementia health-related quality of life". In: *Health Technology Assessment* 9.10 (2005), pp. 1–93. DOI: `10.3310/hta9100`.

[16] Gunther Eysenbach. "The law of attrition". In: *Journal of medical Internet research* 7.1 (2005), e402.

[17] Amit Baumel et al. "Objective user engagement with mental health apps: systematic search and panel-based usage analysis". In: *Journal of medical Internet research* 21.9 (2019), e14567.

[18] Barry Reisberg, Ramu Vadukapuram, and Sunnie Kenowsky. "The global deterioration scale (GDS)". In: *World Alzheimer Report* 2022 (2022), p. 44.

[19] Frank Wilcoxon. "Individual comparisons by ranking methods". In: *Biometrics Bulletin* 1.6 (1945), pp. 80–83. DOI: `10.2307/3001968`.

[20] Jacob Cohen. *Statistical Power Analysis for the Behavioral Sciences*. 2nd. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

[21] Norman Cliff. "Dominance statistics: Ordinal analyses to answer ordinal questions". In: *Psychological Bulletin* 114.3 (1993), pp. 494–509. DOI: `10.1037/0033-2909.114.3.494`.

[22] Jeanine Romano et al. "Appropriate statistics for ordinal level data: Should we really be using t-test and Cohen's d for evaluating group differences on the NSSE and other surveys?" In: *Annual Meeting of the Florida Association of Institutional Research*. Cocoa Beach, FL, 2006.

A Weekly Improvement Progression Plots

Weekly (non-cumulative) improvement progression plots for each domain, showing week-by-week caregiver-assessed improvement scores with 95% confidence intervals.

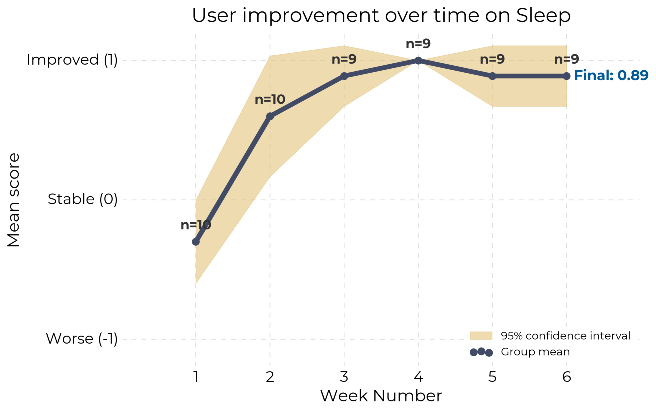

Figure A.1: Weekly caregiver-assessed improvement progression on Sleep

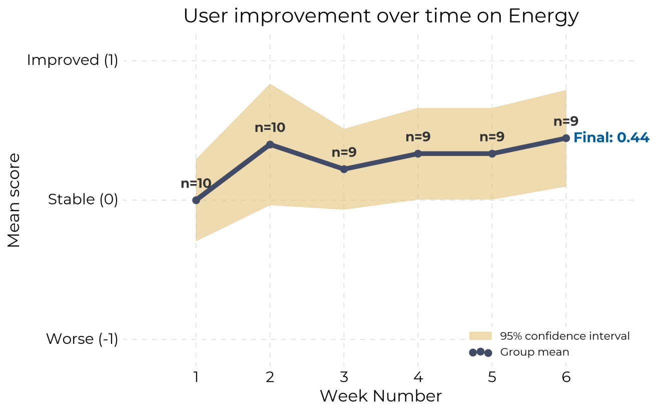

Figure A.4: Weekly caregiver-assessed improvement progression on Energy

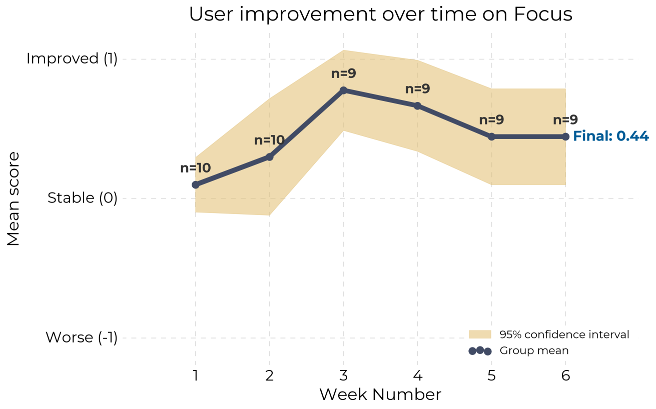

Figure A.2: Weekly caregiver-assessed improvement progression on Focus

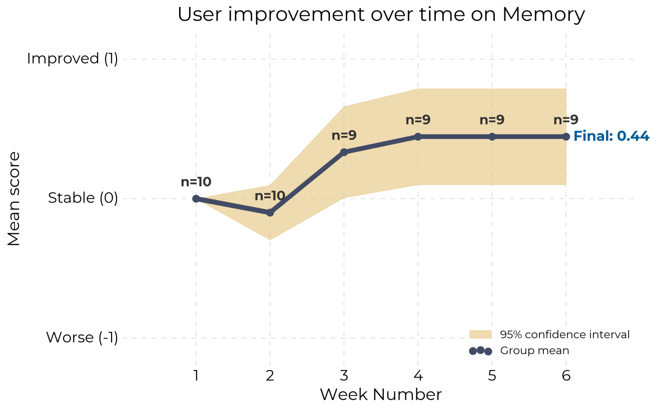

Figure A.5: Weekly caregiver-assessed improvement progression on Memory

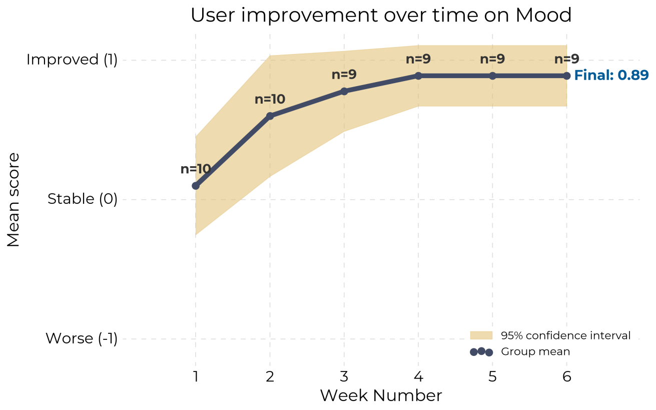

Figure A.3: Weekly caregiver-assessed improvement progression on Mood

B Cumulative Progression Plots with Individual Trajectories

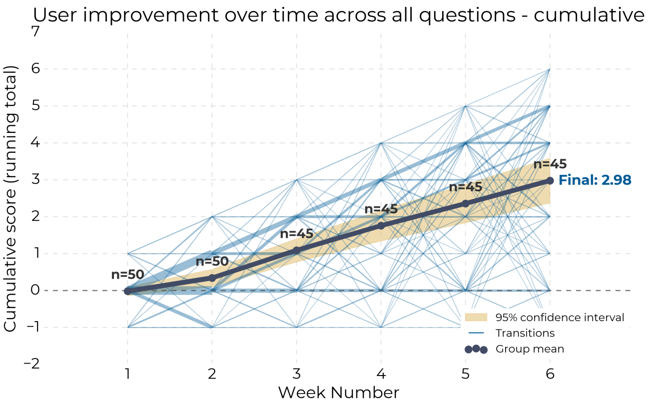

Figure B.1: Cumulative improvement with individual trajectories – All questions

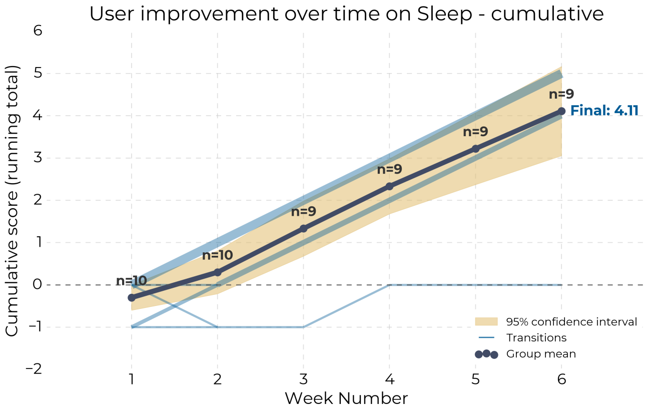

Figure B.2: Cumulative improvement with individual trajectories – Sleep

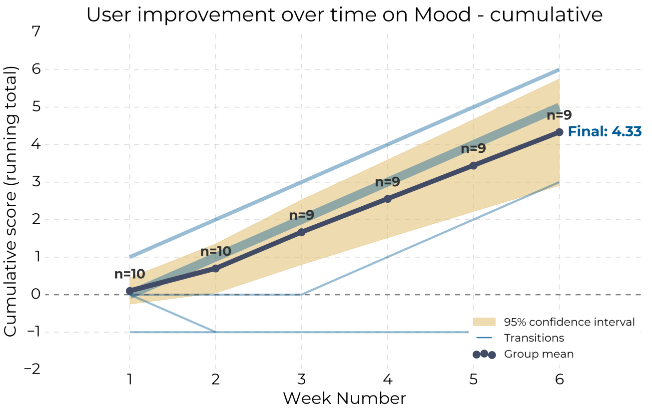

Figure B.3: Cumulative improvement with individual trajectories – Mood

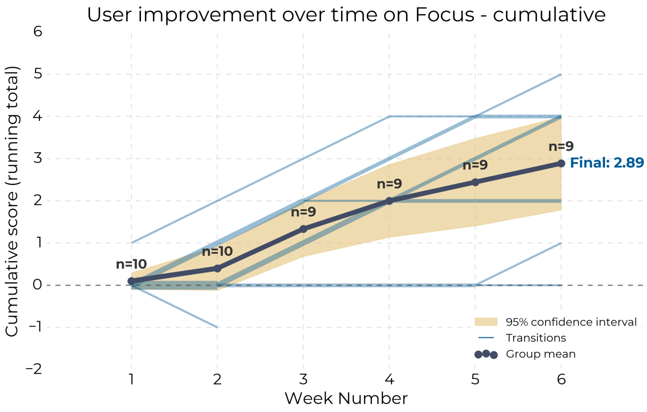

Figure B.4: Cumulative improvement with individual trajectories – Focus

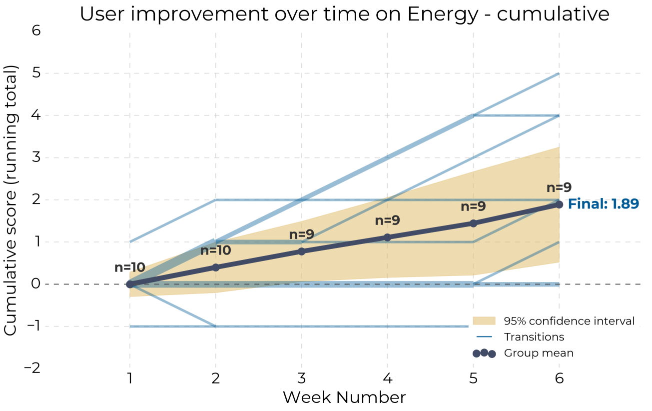

Figure B.5: Cumulative improvement with individual trajectories – Energy

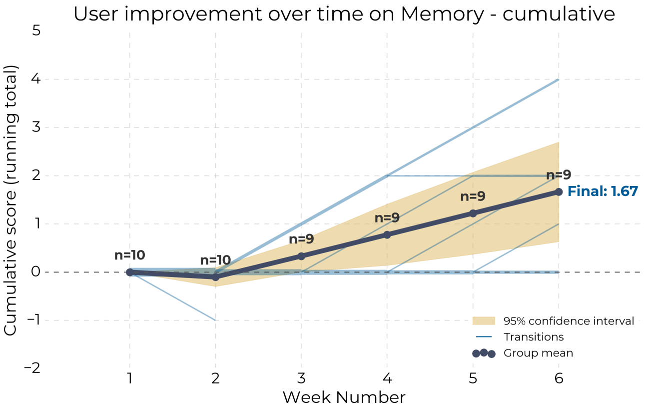

Figure B.6: Cumulative improvement with individual trajectories – Memory

C Absolute Improvement Area Plots by Domain

Per-domain absolute improvement area plots showing the difference between "Improved" and "Worse" assessments each week. The shaded area represents the cumulative benefit over time.

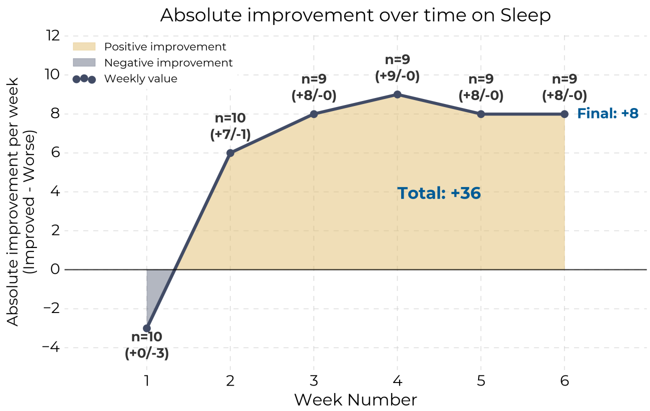

Figure C.1: Absolute improvement area plot - Sleep. Caregivers observed strong and early Sleep improvements.

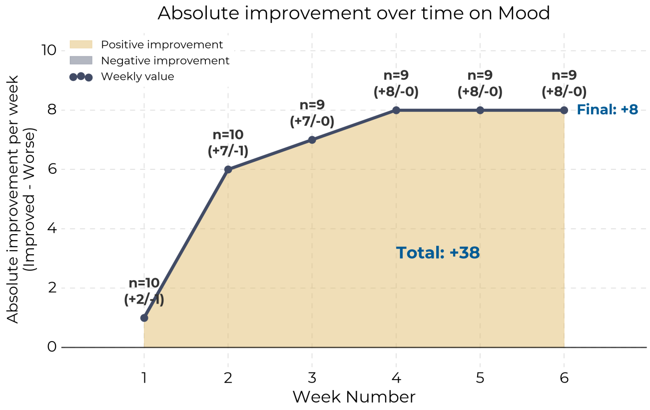

Figure C.2: Absolute improvement area plot - Mood. The highest total improvement across all domains.

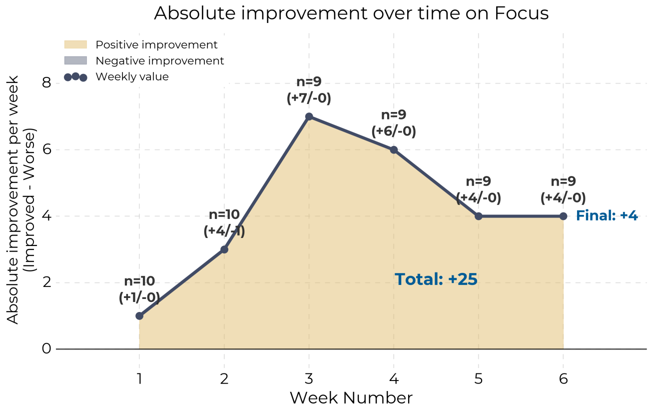

Figure C.3: Absolute improvement area plot - Focus. Consistent positive observations from week 2 onwards.

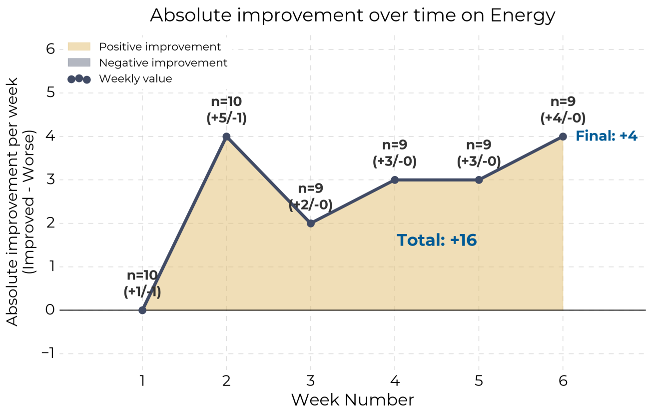

Figure C.4: Absolute improvement area plot - Energy. Moderate but consistent positive observations.

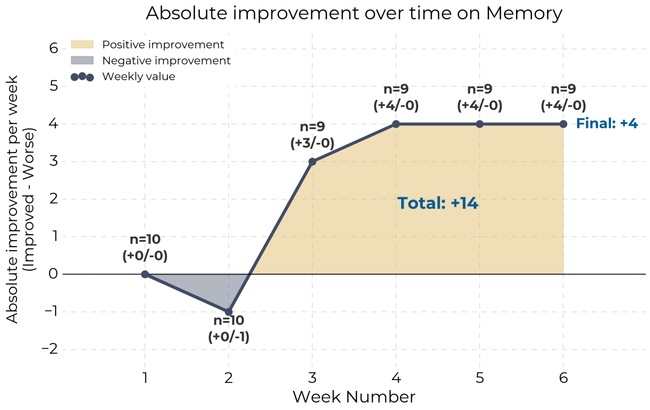

Figure C.5: Absolute improvement area plot - Memory. Note the delayed onset pattern with positive observations emerging from week 3.