

Supplementary Information

for

Multimodal Evaluation of Cognitive Load During Prosthetic Device Use in Activities of Daily Living

Lorena Velásquez, Gabriella Goytizolo, Hasan Ayaz, Jeremy D. Brown

1 Supplementary Instrumentation and Experimental Setup



Figure S1: Sequential procedure for fitting the fNIRS headband and eye-tracking glasses prior to data collection. (A) Initial placement of the fNIRS headband over the participant's forehead. (B) Alignment of the headband relative to anatomical landmarks. (C) Securing and tensioning the headband to ensure stable contact. (D) Placement of an elastic cap to further stabilize optodes during data collection. (E) Placement and adjustment of the eye-tracking glasses.

Before data collection, the fNIRS headband (Model 2000s, fNIR Devices, LLC) was positioned over the participant's prefrontal cortex following standardized placement procedures (Fig. B4). The lower edge of the headband was aligned approximately just above the participant's eyebrows, ensuring coverage of bilateral dorsolateral and medial prefrontal regions. Alignment was confirmed relative to midline anatomical landmarks (glabella and nasion) to maintain consistent optode positioning across participants. Care was taken to smooth down the hair on the head and the eyebrows to prevent any strands from interfering with fNIRS optodes. This was particularly important on the edges of the device, as they aligned with the temples, often the narrowest portion of the forehead. Optodes interacting with unmovable obstacles were noted before further calibration. With these placements, you are tailoring the fit to the person's anatomy, and you aim for the best that you can do. The goal is to preserve as many optodes as possible.

The headband was secured with its integrated Velcro tensioning system and adjusted to ensure firm yet comfortable contact between optodes and the skin. An additional headband was placed on top of the fNIRS band to block ambient light, paying special attention to blocking light on the top of the band, where the optodes meet the skin, and tucking beneath the band to prevent light exposure above the eyebrow. To minimize motion artifacts, the wire leads of the fNIRS optode were fixed behind the participant's head using Velcro. Participants were asked to perform small head movements during setup to confirm stability before moving on to the task. Participants were

also asked about any tightness or pressure points along the band to ensure it was secure but not uncomfortable to wear.

Signal quality was verified and calibrated using FNIRSoft before each experimental block. Channels were inspected for acceptable light intensity levels, minimal saturation, and low ambient light interference. Particular attention was paid to optodes along the superior edge of the band and near the brow line, where loose skin contact or improper shielding could introduce noise. If channels exhibited poor contact, elevated noise, or unstable baselines, the headband tension and ambient light shielding were adjusted, and the optodes were repositioned as needed before proceeding. Data collection began only after stable hemodynamic signals were confirmed across all channels.

Participants were then asked to put on the pupil core glasses as they would any set of glasses. The eye-tracking cameras were then adjusted to capture the participant's pupils. Because the Pupil Core eye-tracking system emits near-infrared light, care was taken to minimize optical interference between systems. The fNIRS headband was positioned above the eye-tracking frame, with the additional shielding used to block out light. During setup, real-time signal inspection would capture and attempt to correct the effects of eye-tracker-based light contamination in fNIRS channels through adjusting signal tuning gains in fNIRSoft. This sequential fitting and validation procedure ensured stable multimodal acquisition throughout task performance. Eye-tracking and fNIRS calibration were performed with the participant standing, head facing down toward the task table to maintain consistency with the head position used during the experiment.

Channels with unstable levels of light contamination were noted. Data was collected from all optodes, but optodes with more than 20% data corruption due to unstable light exposure were excluded from the final dataset. This led to analysis being conducted only on the top row of optodes spanning all four brain regions, as the bottom row was affected by near-infrared light.

2 Supplementary Analysis

2.1 Gaze Allocation

To complement pupillary measures of cognitive load, gaze behavior was analyzed to characterize visual attention strategies during task performance. These analyses were designed to quantify how participants distributed their visual attention across task-relevant elements, including the prosthetic end effector (or natural hand), task objects, and the surrounding workspace. While gaze metrics were not used as a primary measure of cognitive load, they provide important contextual information for interpreting pupillary responses and task-dependent attentional demands.

2.1.1 Data Acquisition and Pre-processing

Eye-tracking data was collected using the Pupil Core system (Pupil Labs GmbH, Berlin, Germany), which records binocular gaze position and pupil diameter at 200 Hz. Raw gaze data were exported and processed offline in the Pupil systems' associated processing software (PupilPlayer). Samples with confidence values below 0.6 were excluded to mitigate the effects of blinks, occlusions, and tracking loss. The remaining data were temporally aligned with task segments corresponding to individual trials.

2.1.2 Regions of Interest (ROI) Definition

To quantify gaze allocation, the visual scene was segmented into task-relevant regions of interest (ROIs). ROIs were defined based on consistent spatial features of the experimental setup and included:

- **End Effector:** the participant's hand (natural or prosthetic)
- **Task Objects:** objects actively manipulated during each task (e.g., blocks, checkers, cartons)
- **Target Area:** destination locations for object placement or transfer
- **Other:** all remaining areas within the scene not directly involved in manipulation

ROI boundaries were defined in the image plane of the scene camera and were applied consistently across trials. Specific ROI labels for each task are provided in the corresponding supplementary figures.

2.1.3 Gaze Allocation Metrics

Gaze behavior was quantified as the proportion of fixation events directed toward each predefined region of interest (ROI) during a trial. Fixations were identified using a dispersion-based algorithm with a threshold of 1.5° and a duration window of 100–4000 ms, selected to capture both brief and sustained fixations while maintaining spatial precision in a visually complex environment. For each trial, ROI labels were manually assigned to fixation events based on their spatial location

within the scene. Gaze allocation was then computed as the proportion of total fixation events directed toward each ROI, obtained by normalizing the number of fixations within each region by the total fixation count for that trial. These proportions were averaged across trials within each task and condition. This metric provides a summary of visual attention distribution across task-relevant elements without relying on fixation duration, which can be sensitive to task dynamics and environmental variability.

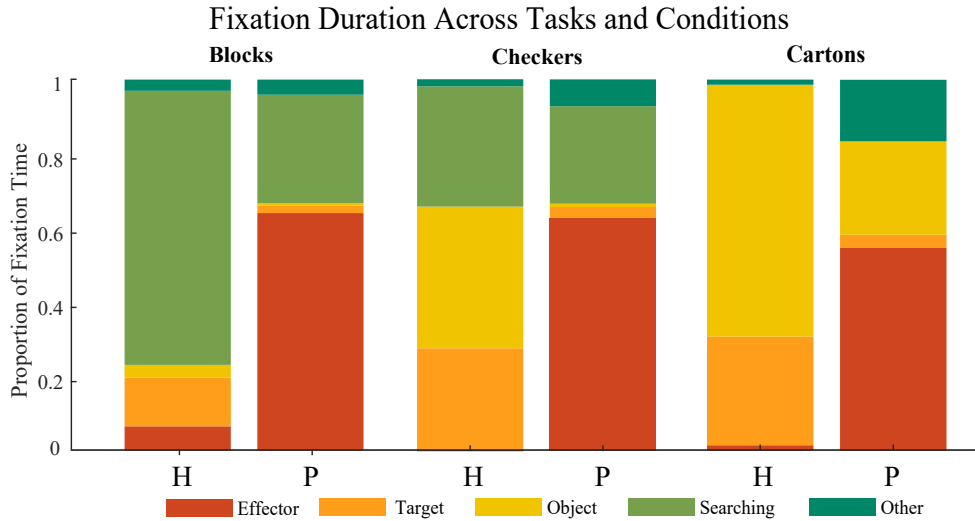


Figure S2: Proportion of fixation events directed toward predefined regions of interest (ROIs), including the end effector, target area, task objects, searching, and other regions, during the block transfer, checker transfer and stack, and carton transfer and pour tasks across both end effector conditions, natural hand (H) and prosthesis (P). Values represent the mean proportion of fixation events averaged across trials and participants for each task and condition.

2.2 Exploratory Multimodal Covariate Analysis

To further explore potential interactions between workload modalities, exploratory covariate analysis was conducted between the two biologically derived measures collected in this study: pupillary dynamics (LHIPA) and prefrontal hemodynamic response (Δ HbO). This comparison was motivated by the alignment observed between fNIRS and subjective workload measures, while eye-tracking metrics demonstrated more task-dependent divergence.

We evaluated whether pupillary workload metrics explained additional variance in the neural response measured through fNIRS. Representative ROI-specific mixed-effects models were evaluated for the block transfer task, which demonstrated significant condition-related effects in both fNIRS and eye tracking, and the weighted transfer task, which demonstrated significant effects in fNIRS but not eye tracking. Nested models were compared using likelihood ratio tests to assess whether inclusion of LHIPA as a covariate improved model fit for Δ HbO responses. Base models included condition and trial as fixed effects with participant as a random intercept, while covariate models additionally included LHIPA. Inclusion of LHIPA did not significantly improve model fit in either task condition (all $p > 0.69$), and condition-related increases in prefrontal activation remained significant following inclusion of the pupillary metric.

3 Supplementary Results

3.1 Task Metrics

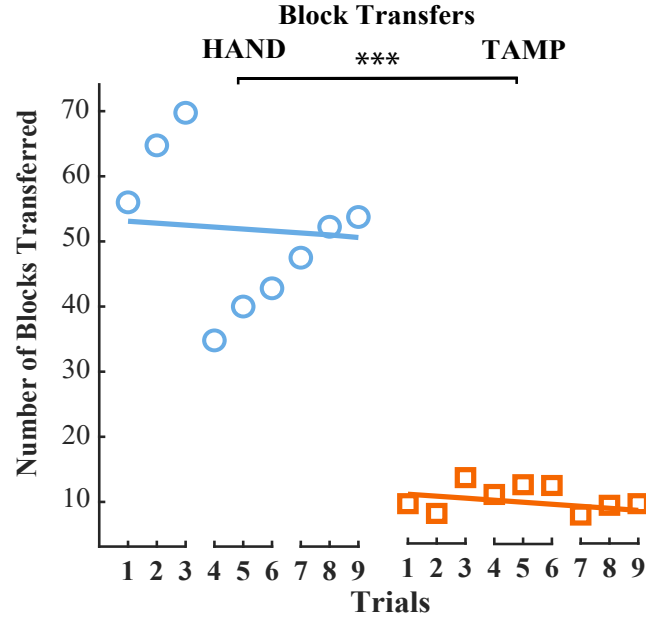


Figure S3: The average number of block transfers for each condition across trials, where the individual data points represent the average for each trial (for all participants in each condition), and the solid lines indicate the model's prediction. * indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$.

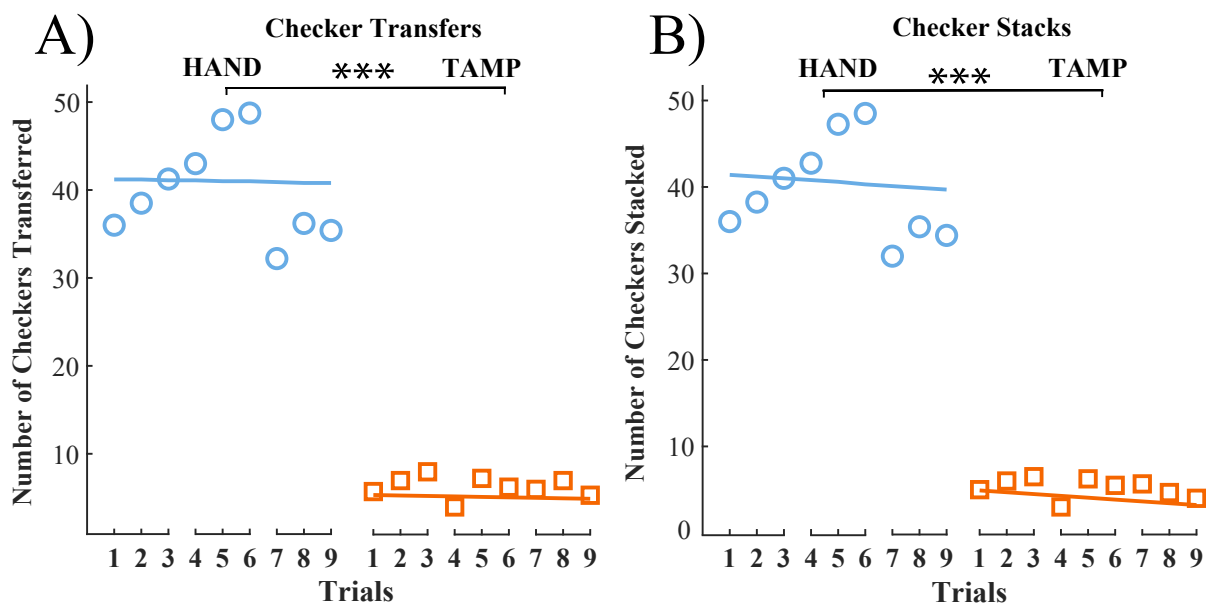


Figure S4: The average number of checker: (A) transfers and (B) stacks for each condition across trials, where the individual data points represent the average for each trial (for all participants in each condition), and the solid lines indicate the model's prediction. * indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$.

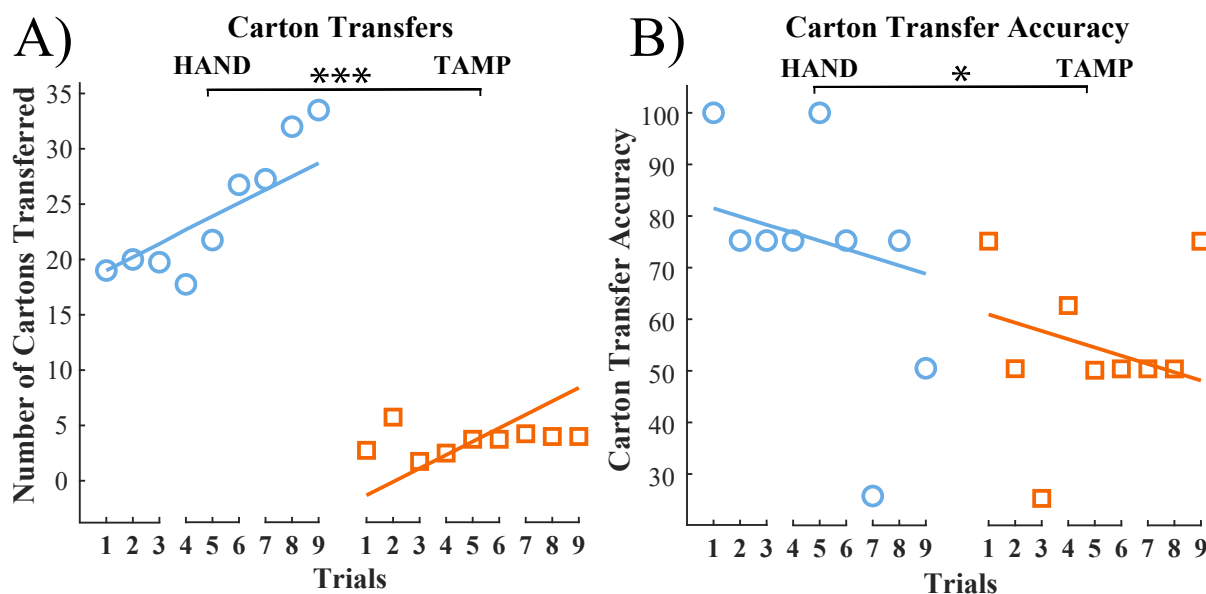


Figure S5: The average number of carton: (A) transfers and (B) transfer accuracy for each condition across trials, where the individual data points represent the average for each trial (for all participants in each condition), and the solid lines indicate the model's prediction. * indicates $p < 0.05$, ** indicates $p < 0.01$, and *** indicates $p < 0.001$.