DATA AND SAFETY MONITORING PLAN

Data and safety monitoring will be performed by the applicant, under the guidance of Prof. Matthis. All subject identifiers will be removed from the data. Electronic data will be maintained in digital format stored without subject identifiers and only with the individual study ID number. Copies of the data will be stored in a locked cabinet and on a secure large storage capacity data server, through which only authorized personnel will be permitted entry. Only the PI and other approved personnel working with the PI on this study will have access to the de-identified data. All unpublished data will be kept for a minimum of five years. Published data will be hosted by reputable online services that provides an official DOI for each item (e.g. Figshare, Mendeley Data).

Types of Data

The primary data generated by this research will consist of digitally sampled time histories of continuous physical measurements, including motions and forces. Kinematic data will be sampled at 300 Hz (minimum). Preliminary experiments indicate that the primary data generated by a single experimental session may occupy roughly 1GB of storage space (not including video). Videos will be encoded in H.265 (.mp4) format when possible, adapting appropriately as video encoding standards progress.

Secondary data generated by this research will consist of details of experimental subjects and the progress of an experiment (date, time, duration, equipment calibration and date, unusual circumstances). With the exception of subject identity as described above, these secondary data will initially be recorded in laboratory notebooks of the responsible investigator. Subsequently they will be transcribed to digital format and associated with the primary data.

Meaningful use of the primary data depends on the associated secondary data. Metadata generated from those sources will include statistical and time-series analyses (primarily performed using Matlab, Python, and/or R). These metadata may take any of several forms, including tables, graphical objects, animations and real-time interactive simulations.

Data Storage and Security

Data generated by the acquisition systems used in this research (specifically motion acquisition) will initially be recorded in native instrument binary format. For analysis, it will subsequently converted to MATLAB compatible “mat” format. After quality control, primary and secondary data will be converted to “ascii” format for long-term storage. We have chosen “ascii” format as it is the most long-lived and stable data format known to the investigators.

Metadata will be stored in native binary format and, where practical, will also be converted to “ascii” format for storage. However, this is undesirable and infeasible for some classes of metadata, including animations and real-time simulations.

All data generated will initially be stored on lab computers in the participating institutions. Short-term data backup will be performed at least weekly. Long-term backup of all data will be performed using Northeastern’s secure online backup service. Access to long-term data will be by using standard password procedures.

Data Sharing

Data will be made publicly available only after they have passed stringent internal quality control procedures. Quality control will be based on the extensive prior experience of Professor Matthis. The primary purpose of quality control is to: (i) ensure internal consistency of each dataset; (ii) eliminate the likelihood of data corruption due to equipment malfunction or unacceptable deviation from experimental procedure; and (iii) ensure the appropriateness and validity of any meta-analyses. An essential element of quality control will be review of data by presentation to both research groups. We therefore anticipate that a period of up to six months may be required for satisfactory quality control.

We anticipate that the data captured and created by the proposed research will be of broad interest to communities engaged in research on human motor behavior and its clinical applications. Once they have passed the quality control procedures outlined above, data generated by this research project will be made publicly accessible by posting to reputable online services that provide DOIs for hosted items.