

**FDA DOCUMENTATION FOR THE GENERIC CLEARANCE
OF REQUEST FOR DATA TO SUPPORT SOCIAL AND BEHAVIORAL RESEARCH
(0910-0847)**

TITLE OF INFORMATION COLLECTION: Collection of Data to Support Training Decay Selection for Medical Product Usability Validation Testing

DESCRIPTION OF THIS SPECIFIC COLLECTION

1. Statement of Need:

The Food and Drug Administration (FDA) requires that medical devices undergo usability validation testing to assess and reduce risks associated with the device use. Intended users undergo a training decay period prior to validation testing to simulate use scenarios and better identify critical use errors. The research aims to improve and streamline critical use error identification when performing usability testing by identifying reliable and replicable training decay behavior. Identifying generalizable training decay curves could standardize the methods for conducting usability testing for medical products and ultimately improve use error identification, while avoiding an undue toll on manufacturer resources and delays in getting life-improving innovative products to patients.

This proposal offers a unique industry-academic partnership that will advance the field of medical device safety assessments performed by manufacturers and the FDA. UserWise, Inc. conducts usability validation testing for medical products and supports FDA submissions of medical products on behalf of manufacturers. This research will evaluate the importance of training decay length in evaluating the use-safety of medical devices.

2. Intended Use of Information:

Data from the participants will be analyzed to assess use errors, close calls, and difficulties with use of the device. Changes in error rates will be used to calculate training decay curves by task types, including psychomotor, task memory, and cognitive tasks.

3. Description of Respondents:

We expect that participants will have minimal prior clinical training and no prior experience with the devices being assessed. We will recruit up to 139 participants for the study. Participants in the study may be grouped for analysis, based on their demographics. For each group, we will ensure a reasonable degree of geographic and demographic diversity, including literacy, race/ethnicity, and age.

4. Date(s) to Be Conducted:

The sponsor will conduct a pilot study between January 27, 2020, and February 3, 2020, and conduct the full study between November 9, 2020, and January 5, 2021.

5. How the Information Is Being Collected:

Each participant will attend a training session and then perform tasks to generate a baseline performance level. Then, the training decay period will begin. Each participant will return for a usability study session in which his or her performance will be observed and recorded.

The training session, baseline performance session, and usability study session may all amount to 2.5 hours of involvement in total (e.g., 1.5 hours for the training and baseline performance session and one hour for the usability study session).

During each observed performance session, one to three study personnel will be present, depending on the needs of the protocol and complexity of the study. Typically, a moderator walks the participant through the session and focuses on asking follow-up questions to better understand the root cause of use-related issues. An observer observes the interaction and records all data related to behaviors, follow-up questions from the participants, and participant-reported root causes.

6. Confidentiality of Respondents:

Personal details of each participant (i.e., name and address) will be recorded solely for the purpose of giving informed consent, payment, and collecting initial screening information, but from the point at which the testing is initiated (whichever comes first), a numbering system will be used to protect the identity of the participant and to remove the link between personal details and test data.

At no stage during the recruitment process will the recruiters seek to gain access to any medical records for any participant. All participants will be required to sign a consent form prior to the start of the usability study session. The study facilitator will also ensure the participant agreement is completed prior to the start of the participant's first study session.

Personal details will be kept secure by the study team and will not be circulated to any third party.

All study personnel interacting with human subjects as part of this research will have completed training on the protection of human subjects.

7. Amount and Justification for Any Proposed Incentive:

Participants will receive a token of appreciation to compensate them for their opinions. The payments were calculated using current industry best practice and are kept as low as reasonably possible in order to avoid undue pressure to participate. Participants will receive a token of appreciation for participating in the full study. The total token of appreciation for participating in the full study will be \$115, comprised of:

- \$75 for the training session (first 90-minute session); and
- \$40 for the usability session (60-minute session).

This research study requires participants to return after a period of time to assess the level of training decay that may have occurred. The value of the compensation for this group was determined with the consideration that participants must return for the usability session so that we may assess how much their skills have decreased over time. For example, if the compensation is not adequate, participants may elect to participate in the training session and not return for the usability session. Low participation may result in inadequate data collection or loss of government funds associated with moderator and observer time. Additionally, low participation can cause delays in launching the research, both of which may lead to increased cost.

8. Questions of a Sensitive Nature:

We do not expect any questions that would be sensitive in nature. Participants will be allowed to abstain from answering if they are uncomfortable with any questions asked during the usability study. Additionally, participants will be notified at the beginning of the study session that they are welcome to end the session and leave at any time for any reason.

9. Description of Statistical Methods:

Use errors and difficulties will be compared between each participant's baseline session and the second session to help quantify the level of retained knowledge and allow for comparison between lengths of training decay and participant performance. In addition, use errors and difficulties will be assessed and compared between task types to inform the methodology for the summative usability study. Further statistical analysis will be performed on the study data collected, as described below.

Data will be entered in a linear mixed effects model, an extension of the regular linear model that allows for a nested structure in the data. For this study, the two nested grouping levels will be participant and timepoint.

Each participant will have two observations: one for his or her baseline assessment and one for his or her follow-up assessment. As such, random intercepts will be specified for the participant grouping variable. This approach leads to a more powerful design as all other effects are examined only after accounting for the within-participant variance.

Timepoint will be entered as a binary variable (baseline versus follow-up) as well as decay length (e.g., one hour versus one day) and task type as categorical variables.

If deemed appropriate, other collected data, such as demographics or baseline cognitive measures, will be entered as fixed effects as well. The inclusion of other parameters, such as random slopes, will be determined by assessing if a model with these parameters

significantly outperforms a model without them, using tests such as the likelihood ratio test.

The dependent variable of interest can be quantified in two ways. One will be the percentage of task success rate given cohort (e.g., 4/10 participants successfully located the correct injection site). When quantified as a continuous variable, the statistical model will allow us not only to detect significant differences but also to arrive at an estimate of those differences. For example, this approach will allow us to quantify how much decay occurs between the first and second timepoint or how much more decay there is for one device type versus another.

The data will also be analyzed more categorically by grading each task completed as a “success,” “failure,” or “difficulty.” This method is more in line with the usual reporting of usability studies and will allow us to apply different levels of rigor when quantifying decay curves.

Testing a two-way interaction between timepoint and decay length will address our first goal of quantifying decay curves. It is hypothesized that this interaction will be significant, as the effect of the timepoint variable (i.e., the training decay between their baseline and follow-up time points) will be different, depending on the length of their decay. Planned post-hoc pairwise comparisons between timepoint groups will allow us to characterize the decay curve.

Statistically significant differences between adjacent timepoint groups (i.e., one hour versus three days) will indicate that significant decay has occurred. Failure to reach statistical significance can mean different things in different contexts. A lack of significance between the baseline group and the one-hour group may indicate that training decay has not occurred yet, while a lack of significance between groups (e.g., three days versus one week) may be interpreted as a “bottoming-out” of the training decay curve.

Testing a three-way interaction between timepoint, decay length, and the task will allow us to address our second goal of examining differences between task types. The previously mentioned two-way interaction between timepoint and decay length represents the training decay curve. An additional interaction between this curve and the task type is expected; that is, the relationship between decay length and timepoint differences depends on the type of task. The between-group differences will be examined in the study report with post-hoc comparisons, looking at the differences between task types at each time point. The unique training curve for each task type will be quantified by performing the same pairwise comparisons between timepoint groups and each task. Resulting data will be assessed to determine if there are correlations between tasks that are similar in nature (e.g., Fact Memory). For these comparisons, the expectation is that there will be no significant difference in the baseline assessments, suggesting that the task types are equally difficult at first.

One crucial additional analysis is examining the relationship between each task’s difficulty and its training decay curve. It is hypothesized that the difficulty of the task is likely to influence memory retention significantly. To analyze this variable, each

participant will be asked to assess the difficulty of each task, using a Likert scale after completing all simulated use portions of the study. This data will be examined at both a task-level (i.e., do tasks rated as more difficult have different decay curves?) and a participant-level (i.e., do participants who find tasks more difficult have different decay curves?). It is possible that manufacturers could additionally use subjective rating information to inform what training decay curve to expect.

While these mixed effects models will address the primary hypotheses, further analyses can be carried out to examine relationships between all our variables, such as looking at training decay versus age or education.

BURDEN HOUR COMPUTATION: *(Number of responses (X) estimated response or participation time in minutes (/60) = annual burden hours):*

Type/Category of Respondent	No. of Respondents	Participation Time (minutes)	Burden (hours)
Training and Baseline Performance	139	90 minutes	208.5
Usability Study Session	139	60 minutes	139.0
Total	139	150 minutes	347.5

REQUESTED APPROVAL DATE: December 2019.

NAME OF PRA ANALYST & PROGRAM CONTACT:

Ila S. Mizrachi
 Paperwork Reduction Act Staff
Ila.Mizrachi@fda.hhs.gov
 301-796-7726

Gretchen Opper
 Office of Surveillance and Epidemiology
Gretchen.Opper@fda.hhs.gov
 240-402-8339

FDA CENTER: Center for Drug Evaluation and Research