

This is expected to yield: 3 pages for the appendix (Consent form, PLS, Tasks sheet), about 1 page for the Methodology, about 4-5 pages for the Results Write Up.

Research question: What is the nature of the human-computer voice interaction with Alexa Skills in the completion of everyday tasks?

Hypothesis: There will be a significant difference between the use of Cooking Buddy, Cooking Buddy PRO and Food2Fork website.

Participants: No requirements

Sample: English speakers. Diverse nationalities. Aged, say, 18-30 ish (of legal age and active technology consumers).

Design: within subjects [Note: I am not worried about learning of practice, as in the real world, the users will also get to practice multiple times, every time they use the skill. Will use randomization for the order of skills and tasks. Prefer that compared to mixed or between-subjects design, primarily in order to compensate for a limited sample size. Please feel free to correct me.]

IV: 1) Skill (Cooking Buddy vs. Cooking Buddy Pro vs. Food2Fork website)
2) Task (Step by Step Guide vs. Target Only)

DV: Error percentage || Duration of completion {ANOVA}

Control Variables: Distractions

TASKS:

- 1) **Step by Step Guide:** Say: <invocation name> to begin the conversation.
(skills only)
Request a recipe with <ingredient> and <ingredient>.
Request the next recipe.
Write down the name of the recipe.
Obtain the ingredients for this recipe.
Hint (for skills only): You can ask Alexa to repeat if you are unsure.
- 2) **Target Only:** Use <enter the name of the skill here> to find the second recipe with <ingredient> and <ingredient>.
Write down the name of the recipe and its ingredients.
Hint (for skills only): You can use <invocation name> to begin your conversation and ask Alexa to Repeat if you are unsure.

Additional data:

- 1) Self-reported rating of the levels of the skill IV. [Likert Scale 5 or 7 points]
{Regression Testing} || [10 points scale rating {Another ANOVA}]
- 2) Self-reported rating of the levels of the task IV. [Likert Scale 5 or 7 points]
{Regression Testing} || [10 points scale rating {t-test}]
- 3) Open ended question(s) about participants experience on completing the tasks, the skills, and recommendations. {Thematic analysis (I guess that will depend on the questions and answers)}

Should I distinguish between people who cook and those who do not?