

Doodle & When2Meet: A User Study

Evangelos Kassos

ekassos@college.harvard.edu

RESEARCH QUESTION AND HYPOTHESES

Need for Cognition

A concept that became extremely relevant in the study was Need for Cognition (NFC), which reflects a person's willingness to undertake difficult mental activities. It is often associated with exploratory behavior and deep learning activities [2].

Hypotheses

First, because When2Meet has some made-for-a-pointer UI elements (click and drag), it can be hypothesized that participants on mobile devices will have a harder time selecting their availability. Related to the our study, it can be hypothesized that participants with higher NFC scores will feel more comfortable using the more complicated When2Meet application. However, one can expect that throughout all groups, Doodle will be perceived as more useful and easy to use. Finally, as described by [3], it can be hypothesized that some participants will display individualist behavior and try to arrange meeting times that better fit them, even if they don't work for some in the rest of the group.

Research Question

Which of the two applications, When2Meet or Doodle, is perceived as more useful and easy to use?

STUDY DESIGN

Need for Cognition

To access their need for cognition, participants were asked to rate how characteristic 18 statements were for them personally by choosing a rating of 1-5 (1: Extremely Uncharacteristic, 2: Somewhat Uncharacteristic, 3: Uncertain 4: Somewhat Characteristic, 5: Extremely Characteristic). The questions are presented in Table 1, and referenced from the widely accepted NFC test presented in [1]. Note that statements 3, 4, 5, 7, 8, 9, 12, 16, and 17 are reverse scored.

Procedures

Participants were asked to complete four different tasks sequentially. The tasks were split into two clusters, the Doodle and When2Meet cluster, based on which scheduling solution they were testing. To ensure uniformness, the order of the

| | |
|-----|---|
| 1 | I would prefer complex to simple problems. |
| 2 | I like to have the responsibility of handling a situation that requires a lot of thinking. |
| 3* | Thinking is not my idea of fun. |
| 4* | I would rather do something that requires little thought than something that is sure to challenge my thinking abilities. |
| 5* | I try to anticipate and avoid situations where there is likely a chance I will have to think in depth about something. |
| 6 | I find satisfaction in deliberating hard and for long hours. |
| 7* | I only think as hard as I have to. |
| 8* | I prefer to think about small, daily projects to long-term ones. |
| 9* | I like tasks that require little thought once I've learned them. |
| 10 | The idea of relying on thought to make my way to the top appeals to me. |
| 11 | I really enjoy a task that involves coming up with new solutions to problems. |
| 12* | Learning new ways to think doesn't excite me very much. |
| 13 | I prefer my life to be filled with puzzles that I must solve. |
| 14 | The notion of thinking abstractly is appealing to me. |
| 15 | I would prefer a task that is intellectual, difficult, and important to one that is somewhat important but does not require much thought. |
| 16* | I feel relief rather than satisfaction after completing a task that required a lot of mental effort. |
| 17* | It's enough for me that something gets the job done; I don't care how or why it works. |
| 18 | I usually end up deliberating about issues even when they do not affect me personally. |

Table 1. Need For Cognition Questions

* Statements 3, 4, 5, 7, 8, 9, 12, 16, and 17 are reverse scored.

tasks within the two clusters was randomized for every participant; then the presentation order of the task clusters was also randomized. We chose not to interject tasks of a clusters in between two tasks of another cluster, so as to minimize the cognitive load caused by the need to switch between the distinct design concepts of the two scheduling solutions. Each task is described in detail below.

When2Meet Availability Filling

To study how users perceive the When2Meet platform and confirm user attitudes as described by [3], participants were asked to complete a When2Meet poll with the goal of scheduling an all-hands meeting with their team. The When2Meet poll contained 5 set responses (Figure 1), with some times being suitable for all participants, and some other times being chosen by two or three participants. To ensure consistency, the participants were provided with a pre-filled work calendar schedule and were told to fill their availability in the When2Meet poll.

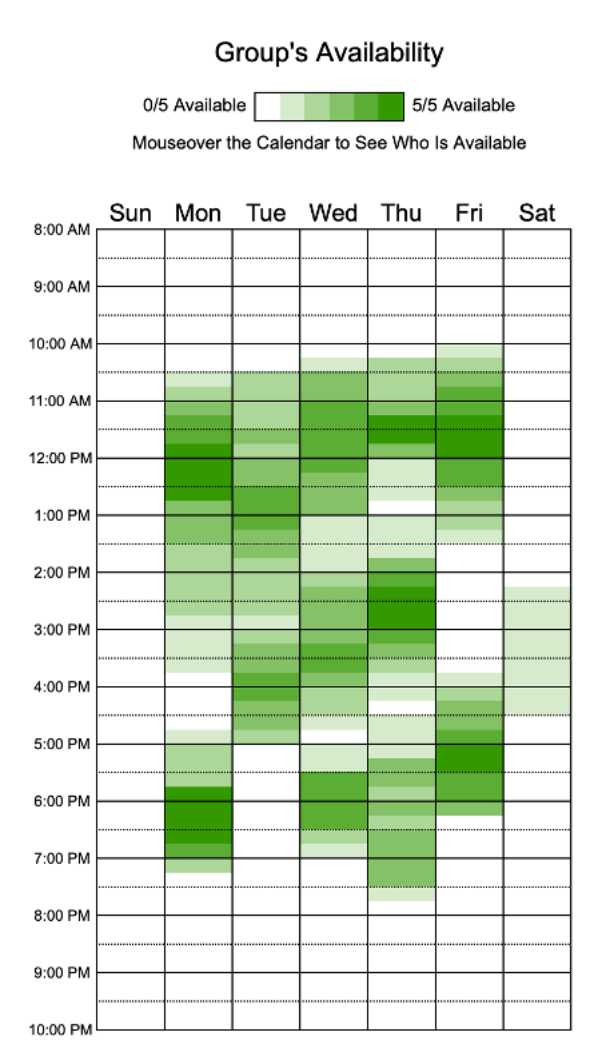


Figure 1. Prefilled When2Meet calendar presented to participants to fill.

Doodle Availability Filling

To enable comparisons between platforms, and once again extend the findings of [3], participants were also asked to

complete a Doodle poll with the goal of scheduling an all-hands meeting with their team. Similar patterns to the When2Meet poll were presented to participants to enable cross-comparisons (i.e., the poll contained 5 set responses, with some times being suitable for all participants, and some other times being chosen by two or three participants). However, the availability windows were shifted so as to prevent participants from cross-referencing their answers between the two tasks. In this task, the “maybe” feature of Doodle was selected for some of the pre-filled responses to test the attitudes of participants towards if-need-be availability, consistent with our third hypothesis.

As described above, the use of When2Meet and Doodle presents two different use cases: filling out one’s availability, and selecting the most suitable time for a meeting based on existing responses. The following two tasks examined participants’ attitudes in the second use case.

When2Meet Meeting Time Decision

To study how efficiently users can make meeting time decisions based on submitted responses by other users, participants were asked to identify the best timeslot for a one hour meeting based on a pre-filled When2Meet poll. They were given a When2Meet poll that contained 5 set responses, and were provided with a pre-filled work calendar schedule they were told to assume was theirs. The When2Meet responses were designed in a way such that the responders had a timeslot that worked well for all users who filled out the poll, but was not ideal for the participant (either between a chain of meetings or very late into the workday, stretching after 5pm). There were a couple of timeslots that were ideal for the participant but would exclude a poll respondent for at least half the meeting. All other possible timeslots were non-optimal for both the participant and the When2Meet respondents.

Doodle Meeting Time Decision

To enable comparisons between platforms, participants were asked to identify the best timeslot for a one hour meeting based on a pre-filled Doodle poll. Similar patterns to the When2Meet poll were presented to participants to enable cross-comparisons (i.e., the poll contained 5 set responses), and participants were provided with a pre-filled work calendar schedule they were told to assume was theirs. However, the availability windows were shifted so as to prevent participants from cross-referencing their answers between the two tasks. The Doodle responses were designed in a way such that the responders had a timeslot that worked well for all users who filled out the poll, but was not ideal for the participant (either between a chain of meetings or very late into the workday, stretching after 5pm). There were a couple of timeslots that were ideal for the participant but would exclude a poll respondent for at least half the meeting. Two timeslots had one respondent’s answer marked as “maybe” to test the attitudes of participants towards if-need-be availability, consistent with our fourth hypothesis. All other possible timeslots were non-optimal for both the participant and the Doodle respondents.

Participants

To understand how useful and efficient the current iterations of the Doodle and When2Meet are for users, a diverse group of

users was invited to participate in a usability evaluation study. The group contained four different subgroups:

- 8 users that had never used the When2Meet application before, but had familiarity with the Doodle application (*When2Meet novice group, Doodle experienced group*),
- 8 users that had never used the Doodle application before, but had familiarity with the When2Meet application (*When2Meet experienced group, Doodle novice group*),
- 8 users that had never used either the When2Meet or Doodle applications before (*When2Meet novice group, Doodle novice group*), and
- 8 users that had used both the When2Meet and Doodle applications before (*When2Meet experienced group, Doodle experienced group*).

The study was limited to current Windows, Android, iOS and macOS users that had been using the respective device for at least six months, to ensure that users were familiar on how to operate said devices. All devices were set up and provided by the researcher to ensure consistency and the study was conducted under the auspices of the researcher to assist with any technical difficulties (none were encountered). Two users from each group used a Windows, Android, iOS and macOS device respectively.

Results

After each task, participants were asked to rate each application on its ease of use and utility on a scale from 1-7, with 1 being “Extremely difficult to use/Would never choose to use again” and 7 being “Extremely easy to use/Would definitely choose to use every time I needed a tool like this.” Therefore, two data points for each application were collected from each participant, for a total of four data points from each participant. The rankings were averaged for each application to obtain the general sentiment of each participant for either application. Table 2 summarizes the results based on the platform the participants completed the study on.

| | When2Meet | | Doodle | | Difference (Doodle-When2Meet) | |
|---------|-------------|---------|-------------|---------|----------------------------------|---------|
| | Ease of Use | Utility | Ease of Use | Utility | Ease of Use | Utility |
| Windows | 5.2 | 4.6 | 6.7 | 6.3 | 1.5 | 1.7 |
| macOS | 5.5 | 5.2 | 7.8 | 7.3 | 2.3 | 2.1 |
| Android | 3.7 | 4.0 | 6.2 | 6.5 | 2.5 | 2.5 |
| iOS | 4.6 | 3.8 | 7.5 | 6.9 | 2.9 | 3.1 |

Table 2. Ratings for Ease of Use and Utility by study participants, grouped by device used.

The results show a clear difference in participants’ perception of utility and ease of use between the two applications, with a strong preference for Doodle. When asked for qualitative feedback (i.e., “why do you feel this way about the two applications?”) participants shared that the Doodle cell-based look resembled a calendar view they were more familiar with and made it easier to navigate. Mobile platform (iOS and Android) participants shared that the When2Meet drag functionality was extremely difficult to navigate, and when they tried to determine the right timeslot for the meeting in the second set of tasks, the interface was even less responsive.

The participants’ responses were also analyzed along the Need for Cognition scale. The participants’ scores were grouped by category (low NFC, moderate NFC, high NFC) and then the average ratings for the participants’ perception of utility and ease of use were calculated for each new group. Table 3 summarizes those results.

| | When2Meet | | Doodle | | Difference (Doodle-When2Meet) | |
|--------------|-------------|---------|-------------|---------|----------------------------------|---------|
| | Ease of Use | Utility | Ease of Use | Utility | Ease of Use | Utility |
| Low NFC | 2.3 | 2.8 | 6.3 | 6.5 | 4 | 3.7 |
| Moderate NFC | 3.6 | 3.9 | 6.9 | 7.2 | 3.3 | 3.3 |
| High NFC | 6.8 | 6.5 | 8.5 | 8.4 | 1.7 | 1.9 |

Table 3. Ratings for Ease of Use and Utility by study participants, grouped by Need for Cognition (NFC).

The results show a clear upwards trend of the participants’ perception of utility and ease of use for both applications. More significantly, the margin between the participants’ perceptions of the two applications shrank with an increase of their Need for Cognition. Once again, however, all groups rated Doodle higher for utility and ease of use.

Finally, the selections of participants for the best meeting time were analyzed and all participants were grouped in terms of whether they chose the best timeslot for all respondents, chose an inconvenient timeslot for some respondents, or chose a non-optimal timeslot for both them and the respondents. The average NFC was calculated and reported in the three categories established above (low NFC, moderate NFC, high NFC). The results are summarized in Table 4.

| | When2Meet | | Doodle | | Difference (Doodle-When2Meet) | |
|---|-------------|----------|-------------|------|----------------------------------|--|
| | Nr. of Par. | NFC | Nr. of Par. | NFC | Nr. of Par. | |
| Chose best timeslot for all | 16 | High | 24 | High | 8 | |
| Chose best timeslot for them; sub-optimal for some | 8 | Moderate | 5 | Low | -3 | |
| Chose non-optimal timeslot for all | 8 | Low | 3 | Low | -5 | |

Table 4. Number of participants and their average Need for Cognition based on the timeslot option that they chose. NFC = Need for Cognition, Nr. of Par. = Number of Participants.

The results show that participants were mostly successful in choosing the best timeslot for both the participant and the poll respondents. However, as also reported by [3], some participants chose the best timeslot for them but a sub-optimal timeslot for some poll respondents. Finally, there were some participants that chose a non-optimal timeslot for both the participant and the poll respondents. Of note is that participants chose a non-optimal/sub-optimal timeslot in the When2Meet task compared to the Doodle task. Moreover, the NFC scores were more widespread when working on the When2Meet task, with participants choosing a suboptimal timeslot having a moderate NFC compared to when working on the Doodle task.

DISCUSSION, FUTURE WORK AND CONCLUSION

The study results offer empirical evidence for all initial hypotheses. First, there was a clear difference in how participants perceived the utility and ease of use of When2Meet and Doodle based on the device they used. As hypothesized, mobile devices provided a worse experience compared to desktop devices. However, more widespread UI elements (checkboxes)

of Doodle provided a better experience across the board compared to the niche UI elements (drag) of When2Meet. Second, participants with a higher tolerance to complexity (higher NFC scores) showed a general willingness to engage with both applications, no matter how complex they are. Nonetheless, as expected, the difference between the two applications in terms of utility and ease of use remained across NFC groups. The study results also verify that there is a group of users who behave in the way that Reinecke et al. reported in [3] by selecting a sub-optimal timeslot for the people they are trying to meet with so that the times works better for them. Finally, the results show that a more complicated application (When2Meet) will make more users choose an incorrect timeslot that is non-optimal for everyone involved. This is also supported by the fact that most participants with high/moderate NFC scores chose the best timeslot for all, while many of those participants ended up choosing an incorrect timeslot (increasing the average NFC scores for each of the groups in the When2Meet task).

These results raise a number of questions for future work both in adapting these applications and developing new meeting scheduling applications. Related to the question of the nature of the device to the participant's perception of utility or ease of use of an application, it is still an open question if these results generalize to other scheduling applications. In the area of NFC, it remains to be seen if the mapping of the simplicity of an application to the types of users that will choose to engage with said application generalizes to more application. Finally,

a larger study could validate these empirical findings and provide statistically significant evidence for the hypotheses examined in this paper.

ACKNOWLEDGMENTS

I would like to personally thank Prof. Elena Glassman, who provided helpful comments throughout the writing process of this project.

DATA AVAILABILITY

The source code can be downloaded from the project's Github repos: <https://github.com/ekassos/when2meet-representation> and <https://github.com/ekassos/doodle-representation>.

REFERENCES

- [1] J. T. Cacioppo, R. E. Petty, and C. F. Kao. 1984. The efficient assessment of need for cognition. *Journal of Personality Assessment* 48 (1984), 306–307.
- [2] A.-M. Cazan and S. E. Indreica. 2014. Need for cognition and approaches to learning among university students. *Procedia-Social and Behavioral Sciences* 127 (2014), 134–138.
- [3] Katharina Reinecke, Minh Khoa Nguyen, Abraham Bernstein, Michael Näf, and Krzysztof Z. Gajos. 2013. Doodle around the world. *Proceedings of the 2013 conference on Computer supported cooperative work - CSCW '13* (2013). DOI: <http://dx.doi.org/10.1145/2441776.2441784>