# ALIGNING EFFECTIVE COMMUNICATION BETWEEN HUMANS AND MACHINES

**Anonymous authors**
Paper under double-blind review

## ABSTRACT

With the growing dependence of humans on technological devices, the influence of digital media on decision-making has become significant. This research focuses on smartphones, which are an always-on, always-with-you medium. An experiment was conducted to understand user interaction with smartphones, and it was found that the time interval between consecutive touches in a browsing session was less than 3 seconds across apps. This finding corroborated the results of previous studies on visual stimuli-driven decision-making scenarios, which suggest that the human brain can decide on an action in a minimum of 0.92 seconds when encountering visual content. Based on these results, this paper proposes a framework, called FSDM (Final Second of Decision Making), for representing digital visual content that is best suited for humans to take the desired action.

## 1 INTRODUCTION

The fifth industrial revolution proposes new challenges, where understanding the synergy between humans and machines becomes most important (Noble et al., 2022). The aspect mainly explored in this paper is that of effective communication for a rapid context-shifting environment. It is assumed that the context, in terms of television, remains constant for about 30 minutes, the average program length. So, the good old 30-second commercial fits within the context of the television medium. In contrast, context-shifting in smartphones is happening at a much faster pace. With every swipe on the smartphone screen, context-shifting is happening. Thus, the visual stimuli for smartphones should be redesigned in a manner that works in less than the context-shifting time on smartphones.

To begin understanding the time required to decide in the brain we take learnings from sports, where the batsman decides which shot to play in 200ms (McLeod, 1987). Looking further we found Facebook (fac) and Netflix studies quoting that the time any user spends considering watching a piece of content is just about 1.8 seconds (abo) in a browsing session. Taking inspiration from these studies, the approach taken by this paper is to understand the minimum and maximum available time for a visual stimulus on a smartphone. The maximum time comes from the experiment as described in the next section. The minimum time comes from studying the processes in the brain and the time it takes to arrive at a decision by looking at a visual stimulus.

## 2 PRIMARY RESEARCH

### 2.1 EXPERIMENT SETUP

The experiment involved 50 volunteers aged 20-30 who downloaded an app to monitor their touches across apps on their smartphones. Data was collected for an average of 2 weeks, with 41 volunteers providing complete data for analysis. The study measured touch intervals, average daily touches, peak touch hours, and total time spent on smartphones. Touch intervals outside the 95th percentile was discarded as non-browsing time, and total time spent on smartphones was taken from another study (Howarth, 2023).

### 2.2 EXPERIMENT RESULT

Smartphone users spend an average of 4 hours and 5 minutes on their device (Howarth, 2023), touching the screen around 5,342 times per day, resulting in a context shift every 2.7 seconds. A constant interaction with the smartphone is observed throughout the day with the highest engagement level occurring at 11 PM. Any persuasive stimuli developed for the smartphone medium should ideally fit within 2.7 seconds, leading to the need for micro-commercials. These findings could potentially bring about a new paradigm for digital communication strategies.

## 3 SECONDARY RESEARCH

Since emotional triggers cue faster responses overall (Carter, 2019), we propose that there needs to be established an emotional trigger at the fore of the visual needs to be played up. Approach or Avoid responses might be cued through establishing Category Need (relevance, context, and significance) and Object Recognition (Desimone & Duncan, 1995; Miller, 2011). We propose that an Emotional Category-Need Trigger, e.g., thirst triggered by looking at visuals of a sandy desert, heat waves, and condensation of water droplets on a chilled bottle of a soft drink. Among all animals, hedonic evaluation of instant experienced utility has been described by Robinson & Berridge (2001) as nonconscious 'Liking'. The activation of these hedonic hotspots indicating 'liking' has been observed when there is recall of past pleasant (hedonic pleasantness) experiences. Such an event might be triggered by Brand Recall and playing up Past Experienced utility of the product. But merely 'liking' does not insinuate action tendencies. Action is motivated through rewards and incentivization, which is described by Robinson & Berridge (2001) as 'Wanting'. We postulate that 'wanting', the nature of which is asserting motivation to perform an action (McClure et al., 2003), could be cued through emotional rewards as well as monetary rewards. The emotional rewards would involve playing up the emotional benefits and emotional highs during use or after using the product. The monetary rewards are played up by emphasizing the price reduction, putting further emphasis on the perception of price reduction as a reward, as opposed to the current method of emphasizing the final, reduced price. This concludes our Final Second of Decision Making (FSDM) framework which would work in just 920 milliseconds (or less than a second) (Braeutigam et al., 2001), which we believe, needs to become the standard practice in digital communication for the 2.7 seconds long interaction time in the rapidly context-shifting smartphone-led world.

## 4 DISCUSSION

We unravel the human-machine interaction that occurs in the decision-making on smartphones by asking the question "What information is the brain looking for?" and 'What representations of information are most relevant?" We approach this problem through the lens of human-machine interaction and the neuroscience of decision-making. Our primary research to understand human-machine interaction quantifies the gaze of a smartphone user to process a piece of content before they make the decision to scroll. We conclude that this time is 2.7 seconds. We further investigate the relevant brain processes that lead to an 'action'. The most confounding discovery about human behavior was that 99.9 percent of all the processing in our brain occurs nonconsciously. Less than 0.5 percent of all available stimuli are brought to our conscious awareness. Our eyes contribute 10,000,000 of the 11,000,000 bits of information to the nervous system, which means they carry the most information out of all our sense organs (Zimmermann, 1989). This approach to guide human-machine interaction can be formalized to understand (1) how the human-machine interaction occurs? (2) what brain processes are involved in driving the desired action? and (3) what constituents drive that behavior? We propose this methodology allows us to learn an appropriate representation of information best suited for the human brain based on learnings of decision-making from neuroscience.

### URM STATEMENT

The authors acknowledge that at least one key author of this work meets the URM criteria of ICLR 2023 Tiny Papers Track.

## REFERENCES

The power of a picture. URL https://about.netflix.com/en/news/the-power-of-a-picture.

Capturing attention in feed: The science behind effective video creative. URL https://www.facebook.com/business/news/insights/capturing-attention-feed-video-creative.

Sven Braeutigam, John F Stins, Steven PR Rose, Stephen J Swithenby, and Tim Ambler. Magnetoencephalographic signals identify stages in real-life decision processes. *Neural Plasticity*, 8(4): 241–254, 2001.

Rita Carter. *The brain book: An illustrated guide to its structure, functions, and disorders*. Dorling Kindersley Ltd, 2019.

Robert Desimone and John Duncan. Neural mechanisms of selective visual attention. *Annual review of neuroscience*, 18(1):193–222, 1995.

Josh Howarth. Time spent using smartphones (2023 statistics), Jan 2023. URL https://explodingtopics.com/blog/smartphone-usage-stats.

Samuel M McClure, Nathaniel D Daw, and P Read Montague. A computational substrate for incentive salience. *Trends in neurosciences*, 26(8):423–428, 2003.

Peter McLeod. Visual reaction time and high-speed ball games. *Perception*, 16(1):49–59, 1987.

Geoffrey Miller. *The mating mind: How sexual choice shaped the evolution of human nature*. Anchor, 2011.

Stephanie M Noble, Martin Mende, Dhruv Grewal, and A Parasuraman. The fifth industrial revolution: How harmonious human–machine collaboration is triggering a retail and service [r] evolution. *Journal of Retailing*, 98(2):199–208, 2022.

Matthew Pelowski, Patrick S Markey, Michael Forster, Gernot Gerger, and Helmut Leder. Move me, astonish me… delight my eyes and brain: The vienna integrated model of top-down and bottom-up processes in art perception (vimap) and corresponding affective, evaluative, and neurophysiological correlates. *Physics of Life Reviews*, 21:80–125, 2017.

Terry E Robinson and Kent C Berridge. Incentive-sensitization and addiction. *Addiction*, 96(1): 103–114, 2001.

Manfred Zimmermann. The nervous system in the context of information theory. *Human physiology*, pp. 166–173, 1989.

## A APPENDIX

A visual stimulus can be perceived, nonconsciously, even when the visual field is not in active pursuit, in as little as 30-100 milliseconds. This process relies on the retinal ganglion cell axon projections to the superior colliculus (SC). A parallel but slower path of axons goes to the lateral geniculate nucleus, the sensory relay station. This is immediately followed by a trace of activation to the pre-frontal cortex (PFC) from the SC in case of emotionally charged visuals. This signal travels from the brain stem to the PFC in less than 1/10th of a second. Through premotor and later motor activation, it signals microsaccades and head movement in the direction of the stimulus. While emotional arousal, interpretation, and awareness, and microsaccades of the eyeballs bring the visual to maximum acuity and full emotional awareness, the visual cortex receives inputs from the thalamus, for interpretation of the visual. The temporal cortices work to interpret and focus the visual. This process is complete in about 350 ms when the PFC can generate a Approach or Avoid response. The next step in the evaluation of a visual is recognition and recall through memory and comparisons with past experiences which seems to occur in the temporal cortices, while the parietal cortices evaluate the significance and meaning of the visual. This process takes about 100 ms (Pelowski et al., 2017). The memory and comparisons then signal and trigger responses from sub-cortical

regions involved in evaluating pleasantness and unpleasantness, consolidation of memories in the hypothalamus, and hypothalamic stimulation via dopaminergic projections that ascribe motivational effects and salience to objects and actions. This aids generation of future of reward predictions and minimization of reward prediction error values for future prediction of actions and expectation of positive reward. The outcomes from these brain regions are visible in about 400 milliseconds from the onset of stimulus (McLeod, 1987), and "Go" motor action is generated in a total of about 920 milliseconds, aggregate as shown in the figure below.

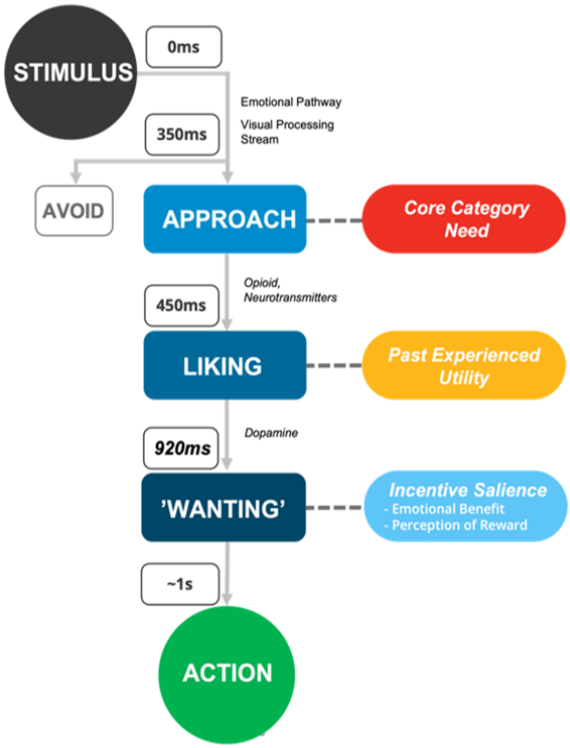


Figure 1: FSDM Framework