

Abstract geometric lines in the top left corner, consisting of several thin, light brown lines that intersect to form various polygons and shapes, creating a modern, minimalist design element.

VPN AND VPS USERS ON CROWDSOURCING PLATFORMS

Ludovic Malfait – ludovic.malfait@bt.com

IMPORTANT NOTE

This work was carried out while the presenter was affiliated with Dolby Laboratories.

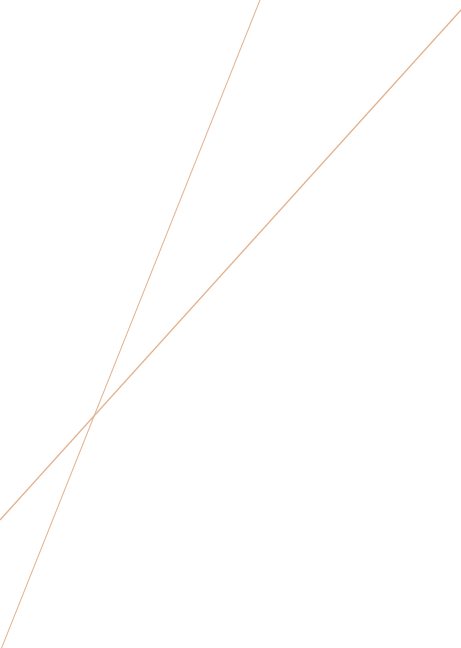
Addressing VPN and VPS users when conducting subjective tests on crowdsourcing platforms.

Malfait, L., Chaudhari, N. & Kim, DS.

Springer Quality and User Experience vol. 10, Art. 2 (2025).

<https://doi.org/10.1007/s41233-025-00072-7>

Direct access: <https://rdcu.be/ehp1t>



CROWDSOURCING PLATFORMS

Pros: convenience

- Wide range of demographics, devices and environments
- Not bound to test facilities
- Fast execution and (can be) cost effective

Cons: lack of control

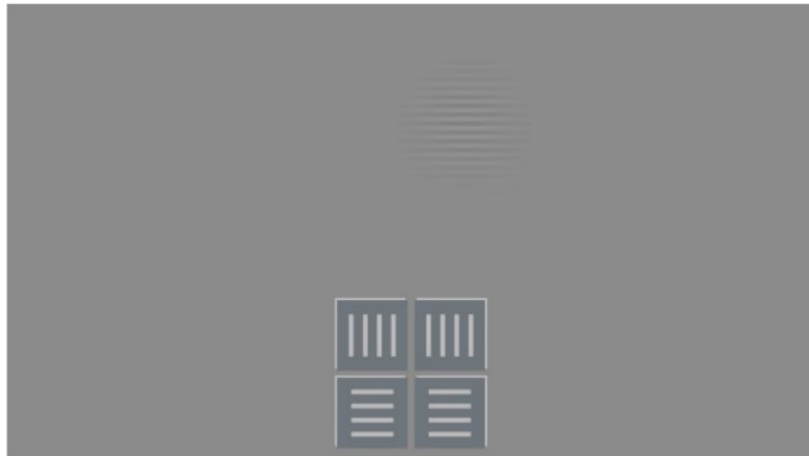
- Participants cannot be monitored
- Tests are typically taken in a web-browser environment, limiting access to device (hardware, operating system configuration)

VIDEO EXPERIMENT

Test to better understand users' sensitivity to image details and resolution on their day-to-day devices

Limited to participants located in the USA (crowdsourcing platform filter)

Designed test software for accurate detection of screen resolution and dot-to-dot rendering of images on the display



(a) Correct presentation



(b) Incorrect presentation

PROBLEM

Initial crowdsourcing test with participants located in the USA:

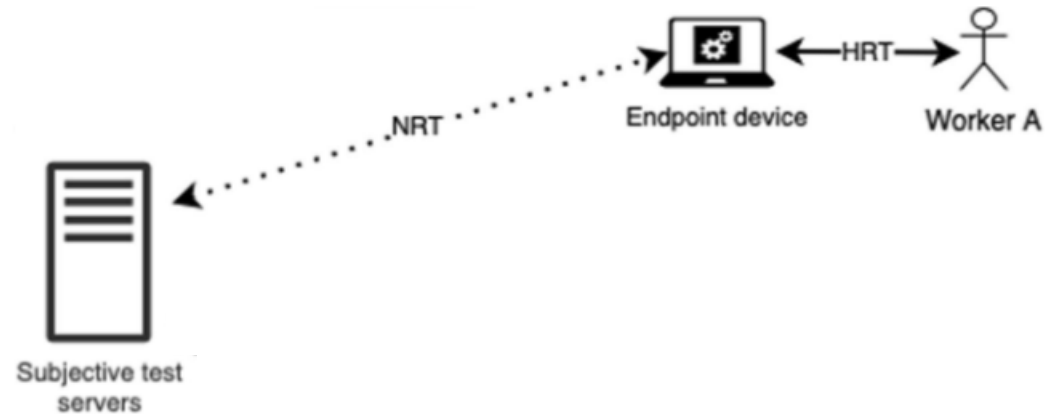
- 55% of the computers had a resolution less or equal to 768p
- Experiment results not aligned with expectation

Computer resolution in the USA from StatCounter:

- 23% of the devices $\leq$ 768p

Investigation revealed that many participants pretend to be in the USA while operating from abroad and masked their location

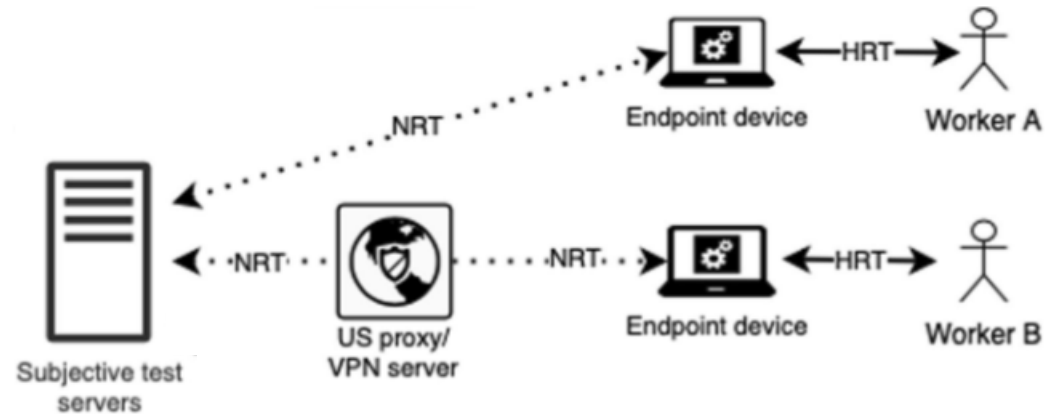
DIFFERENT CONNECTION TYPES



Direct connection

Some non-US IP addresses and non-US time zones

DIFFERENT CONNECTION TYPES



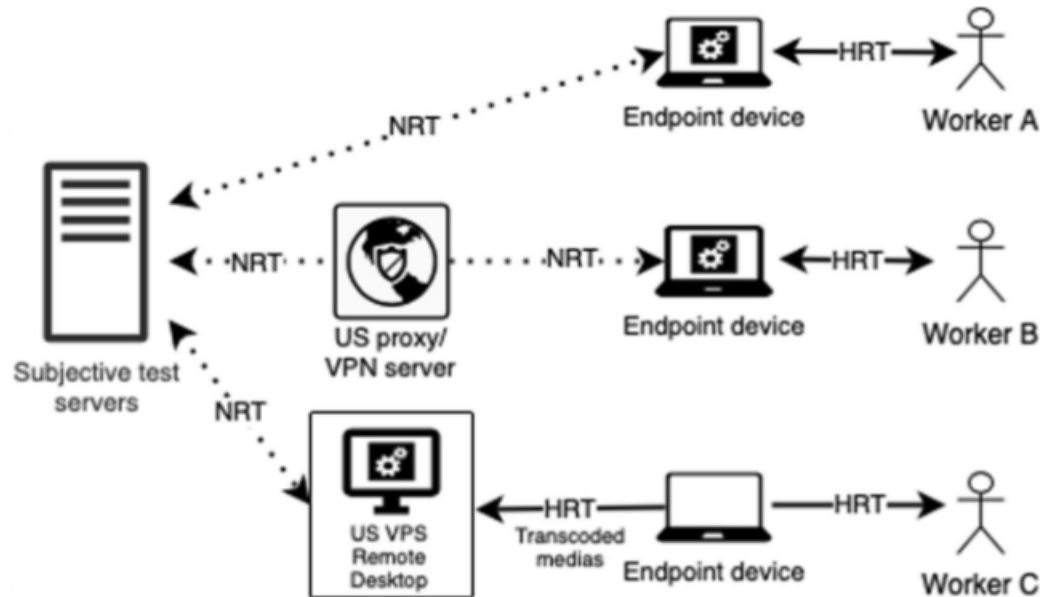
Direct connection

Some non-US IP addresses and non-US time zones

Virtual Private Network (VPN) connection

US IP addresses, some non-US time zones

TYPICAL CONNECTION TYPES FOR CROWDWORKERS



Direct connection

Some non-US IP addresses and non-US time zones

Virtual Private Network (VPN) connection

US IP addresses, some non-US time zones

Virtual Private Server (VPS) connection

US IP addresses and time zones

VPS are a major issue for subjective testing: transcoded contents, virtual audio/video devices, A/V sync, transmission error, RDP client scaling, hard to detect

VPN AND VPS CROWDWORKERS IN THE LITERATURE

Often associated with low quality delivery within the social science research community.

VPN users are typically detected using IP address pools

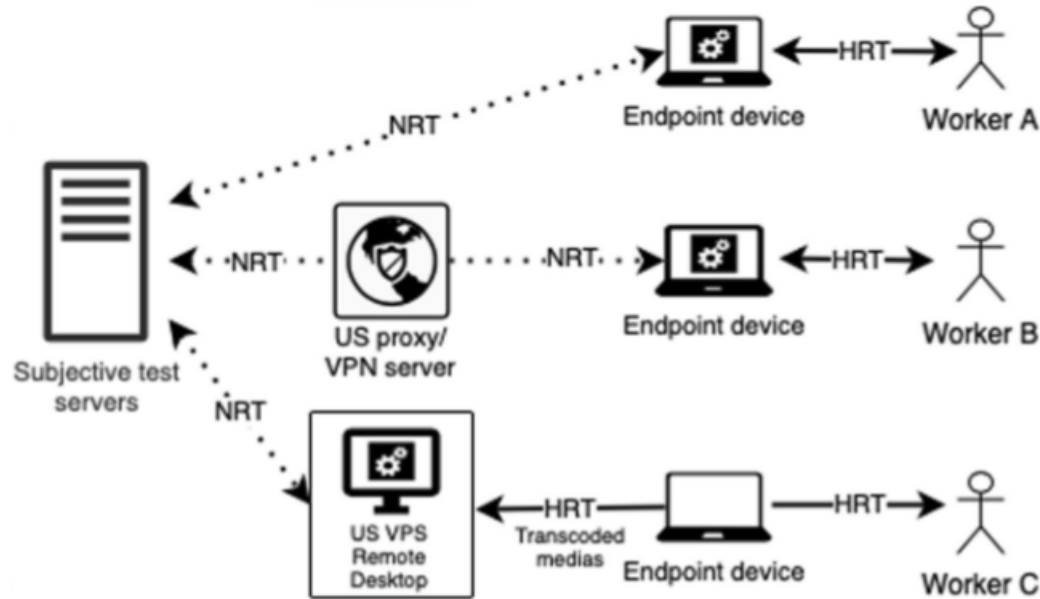
- Constant updates are required
- Could exclude genuine workers

VPS users are typically detected through trapping questions (cultural difference, English comprehension, Human behavioral)

- Not reliable
- Takes away some subjective test time

More objective and reliable detections are needed

ADDRESSING VPN AND VPS USERS



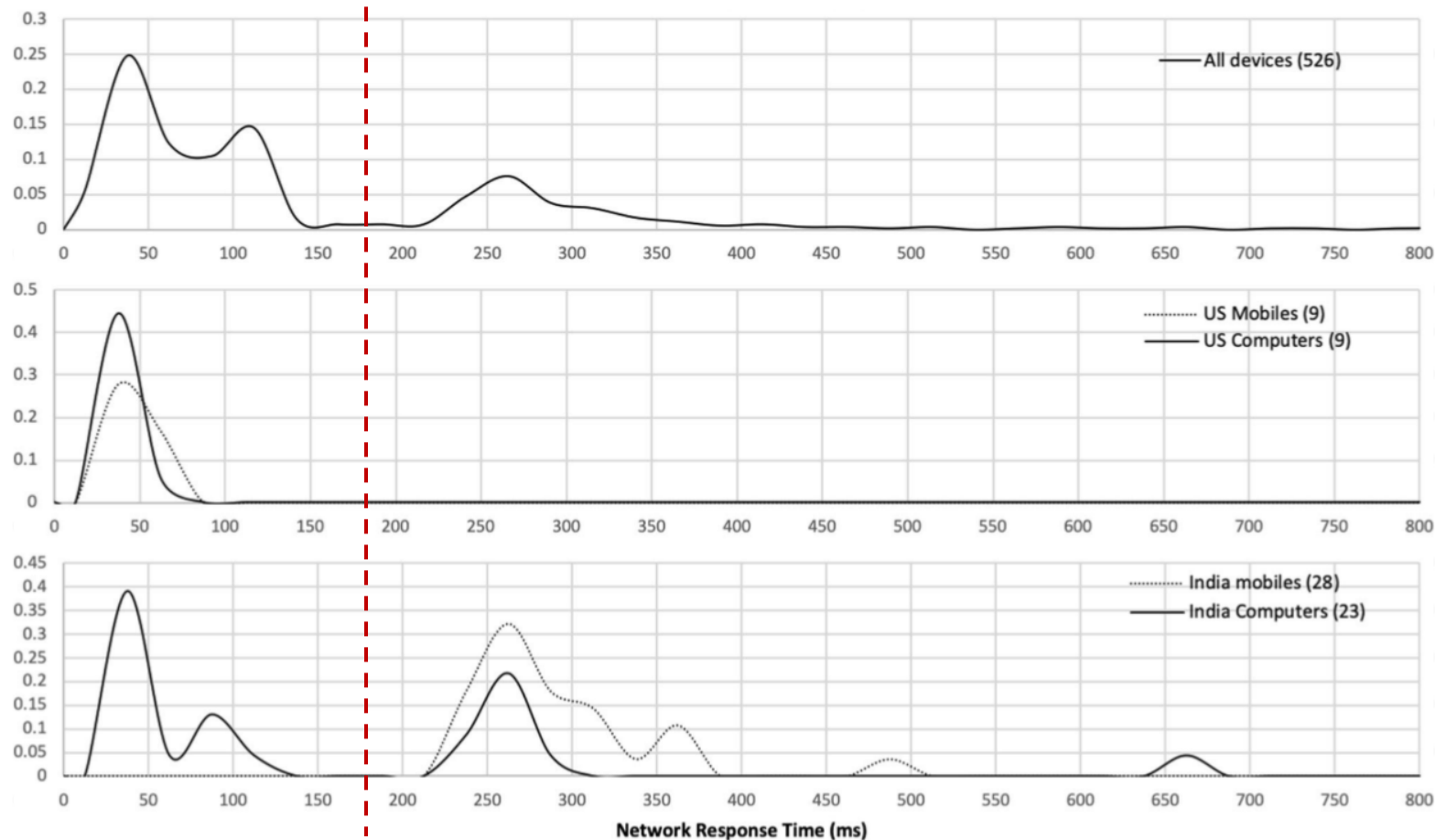
Network Response Time (NRT) – Measure of the transmission time between the subjective test servers and client.

Detects non-US workers using Direct and VPN connections

Human Reaction Time (HRT) – Measure of the time required for a participant to response to a visual stimulus.

Detects non-US workers using RDP clients.

NETWORK RESPONSE TIME MEASUREMENTS



Median latency in the USA is 33 ms for DSL, 63 ms for mobile.

Typical latency between the USA and India is around 215 to 270 ms.

Suggested threshold:
180 ms

18% above the threshold

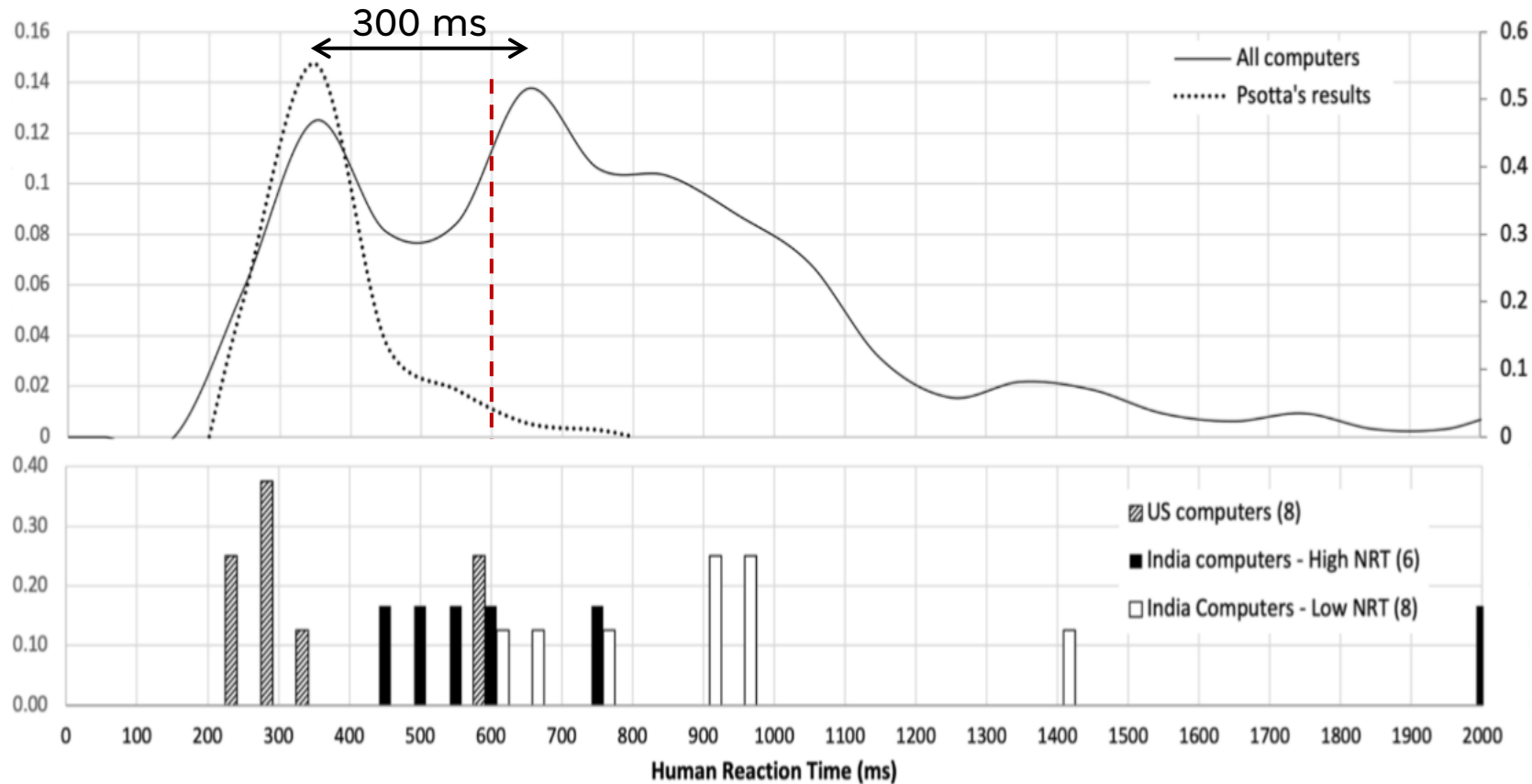
HUMAN REACTION TIME TASK



Simple Reaction Time task (SRT) to a video stimulus.

Participants were asked to press a button as soon as the white rectangle appears on their screen.

HUMAN REACTION TIME MEASUREMENTS

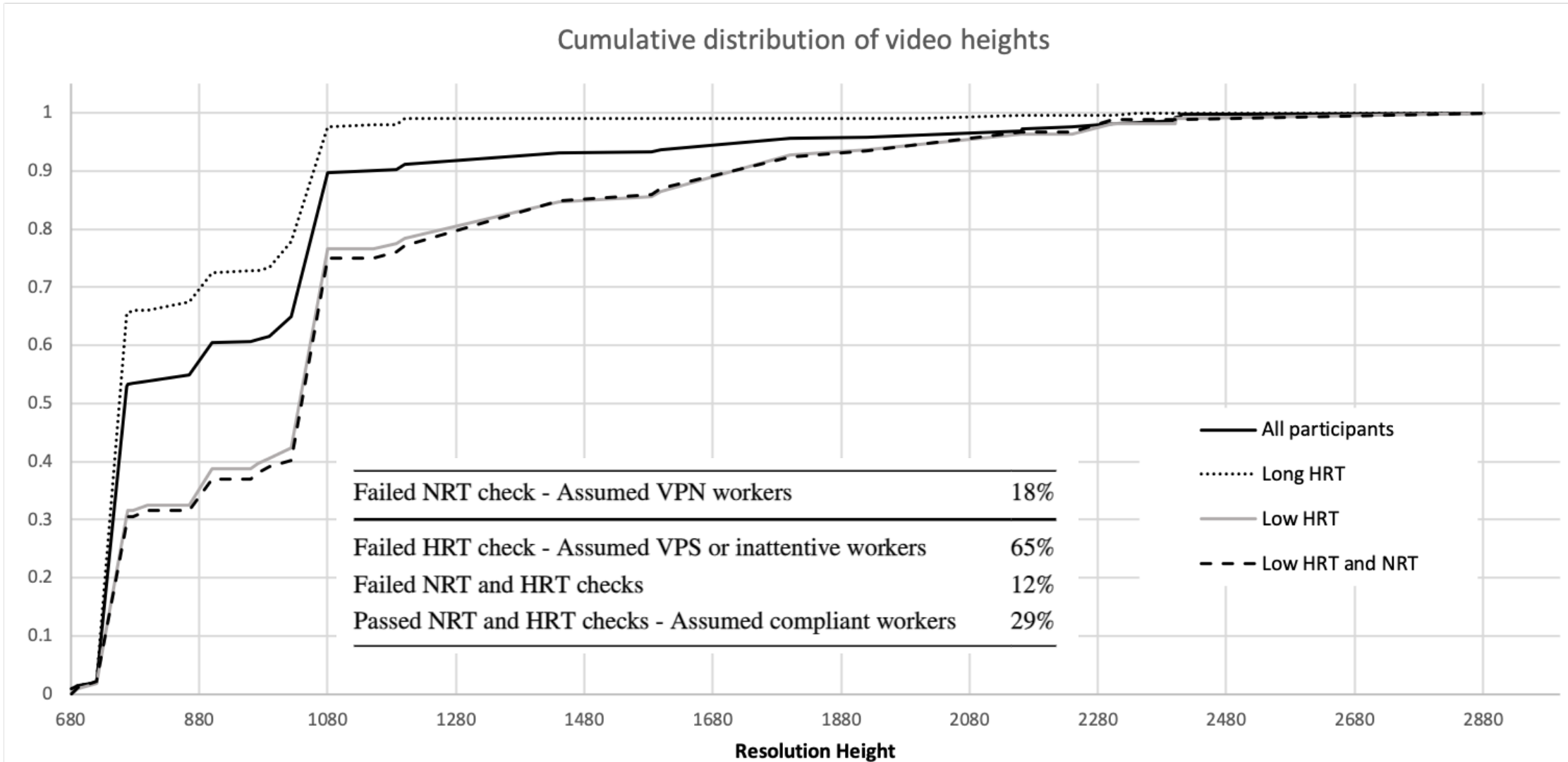


Typical Simple Reaction Time (SRT) to a video stimulus: 356 ms ($\sigma = 80$)

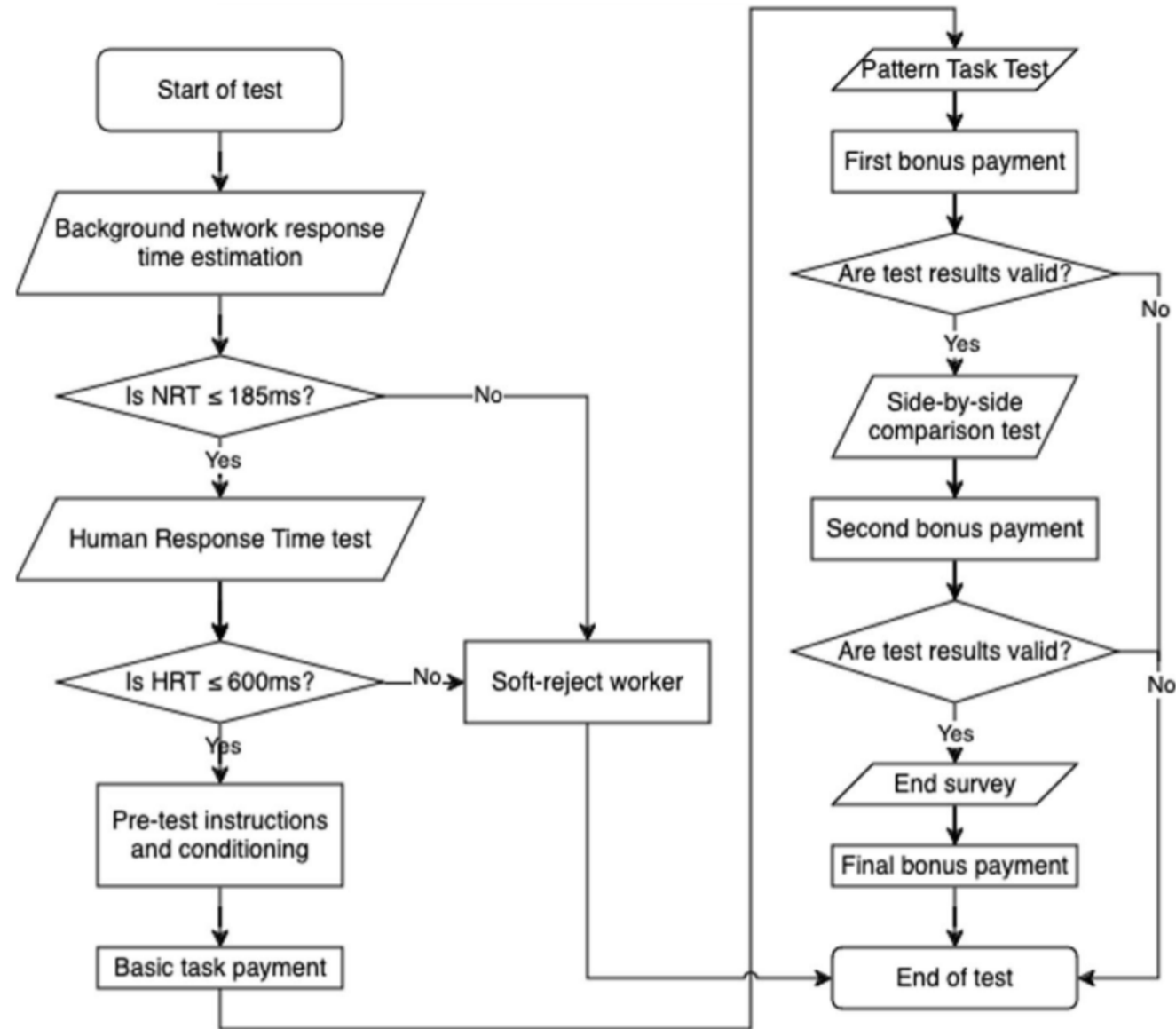
Suggested threshold: 600 ms

65% above the HRT threshold

RELATION BETWEEN HRT AND RESOLUTION



SUBJECTIVE TEST WORKFLOW INTEGRATION, SOFT-REJECTION AND BONUSES



A series of thin, light brown lines forming an abstract geometric pattern in the top left corner of the slide.

SUMMARY

Be cautious when conducting test on crowdsourcing platforms if specific demographics or accurate presentation is required.

Metrics such as NRT or HRT could help collecting more trustable data.

These can be integrated in the test workflow to end the session before starting the test, saving time and money.