The Reverse Turing Test: Individual Differences in the Turing Decision

Vassilis Kallitsis
Deree – The American College of Greece · International Honors Program Thesis · 2025
Supervisor: Dr. Vasileios Vlaseros

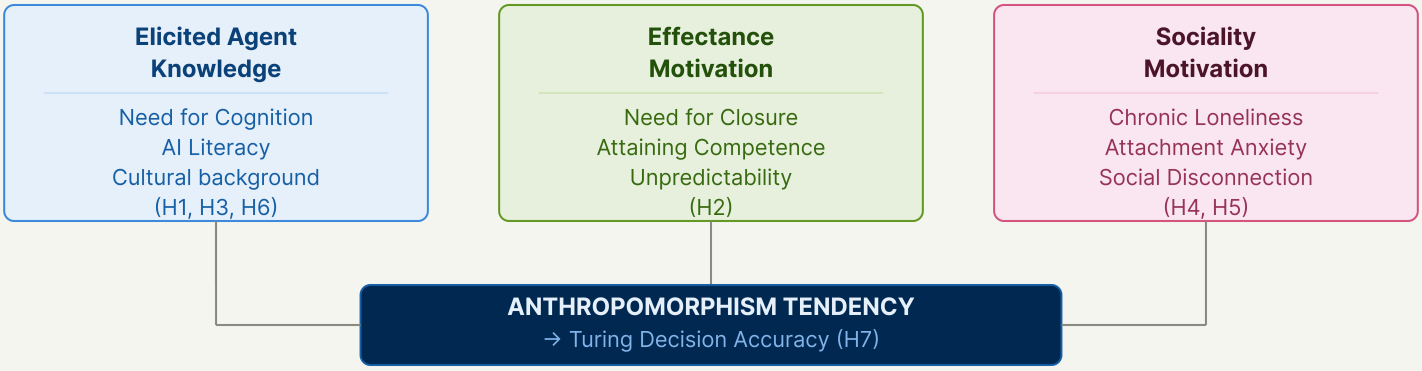
BACKGROUND

As AI-generated text becomes **indistinguishable** from human writing, the ability to identify AI output is critical for avoiding misinformation and preserving genuine communicative action (Habermas, 1984; Giddens, 1984).

The **Turing Decision** — judging whether text is AI or human — is increasingly consequential.

Little is known about which psychological traits predict success. **Anthropomorphism** — attributing human traits to non-human agents — is a prime candidate (Epley et al., 2007).

THEORETICAL MODEL (Epley et al., 2007)



METHOD

Participants

N = 62 undergraduates
Deree / American College of Greece
59.7% female, M age = 22.60
85.5% Greek, 64.5% Psych majors
Convenience sampling, lab setting

Reverse Turing Test Task

50 short texts (~50 words each)
25 human-written, 25 ChatGPT 4.0
Topics: weather, music, sport, TV
Randomised order, lab PCs
Decision: Human or AI?

Psychological Measures

NFC-10 · NFC Scale (15-item)
3-Item Loneliness Scale
ECR Anxiety Subscale
IDAQ · AI Literacy Scale
Multiple linear regression

KEY RESULTS

Human text identification: 63.2% accuracy (M = 15.79 / 25)

AI text identification: 52.6% accuracy (M = 13.15 / 25)

Overall score: 57.9% accuracy (M = 28.94 / 50)

Participants **significantly better at identifying humans than AI:** $t(61) = 35.69$, $p < .001$

Hypothesis	Predictor	Direction	Supported
H1 – Need for Cognition → RTT score		Negative (predicted)	× Not significant
H2 – Need for Closure → RTT score		Negative (predicted)	× Not significant
H3 – AI Literacy → RTT score		**Positive (predicted)**	**✓ Significant**
H4 – Loneliness → AI-text score (negative corr.)		Negative correlation	~ Partial support
H5 – Attachment Anxiety → RTT score		Negative (predicted)	× Not significant
H6 – Ethnic/Cultural Background → RTT score		Differences expected	× Not significant

DISCUSSION & IMPLICATIONS

AI Literacy as Key Predictor

Experience with AI tools predicts both overall and human-text accuracy. Supports AI education interventions to reduce misinformation susceptibility and prevent agency misattribution.

Loneliness & Human Agency

Loneliness negatively correlated with AI-text recognition. Need for social contact may bias individuals toward attributing human agency to AI-generated content.

Keywords: Reverse Turing Test · Anthropomorphism · AI Literacy · Loneliness · Turing Decision · Human-AI Interaction