
# *What Guides Our Choices?* Modeling Developers' Trust and Behavioral Intentions Towards GenAI

**Rudrajit Choudhuri**[1], Bianca Trinkenreich[1], Rahul Pandita[3], Eirini Kalliamvakou[3], Igor Steinmacher[2], Marco Gerosa[2], Christopher Sanchez[1], Anita Sarma[1]



# Generative AI is revolutionizing SE

COMMUNICATIONS OF THE ACM

## The Paradigm Shifts in Artificial Intelligence

Even as we celebrate AI as a technology that will have far-reaching benefits for humanity, trust and alignment remain disconcertingly unaddressed.

## Survey: The AI wave continues to grow on software development teams

TECH

## Google CEO says more than a quarter of the company's new code is created by AI

Hugh Langley Oct 29, 2024, 9:36 PM UTC Share Save

# Trust in AI? *What does it mean? Why does it matter?*

> ***"the attitude that an AI agent will achieve an individual's goals in a situation characterized by uncertainty and vulnerability"***

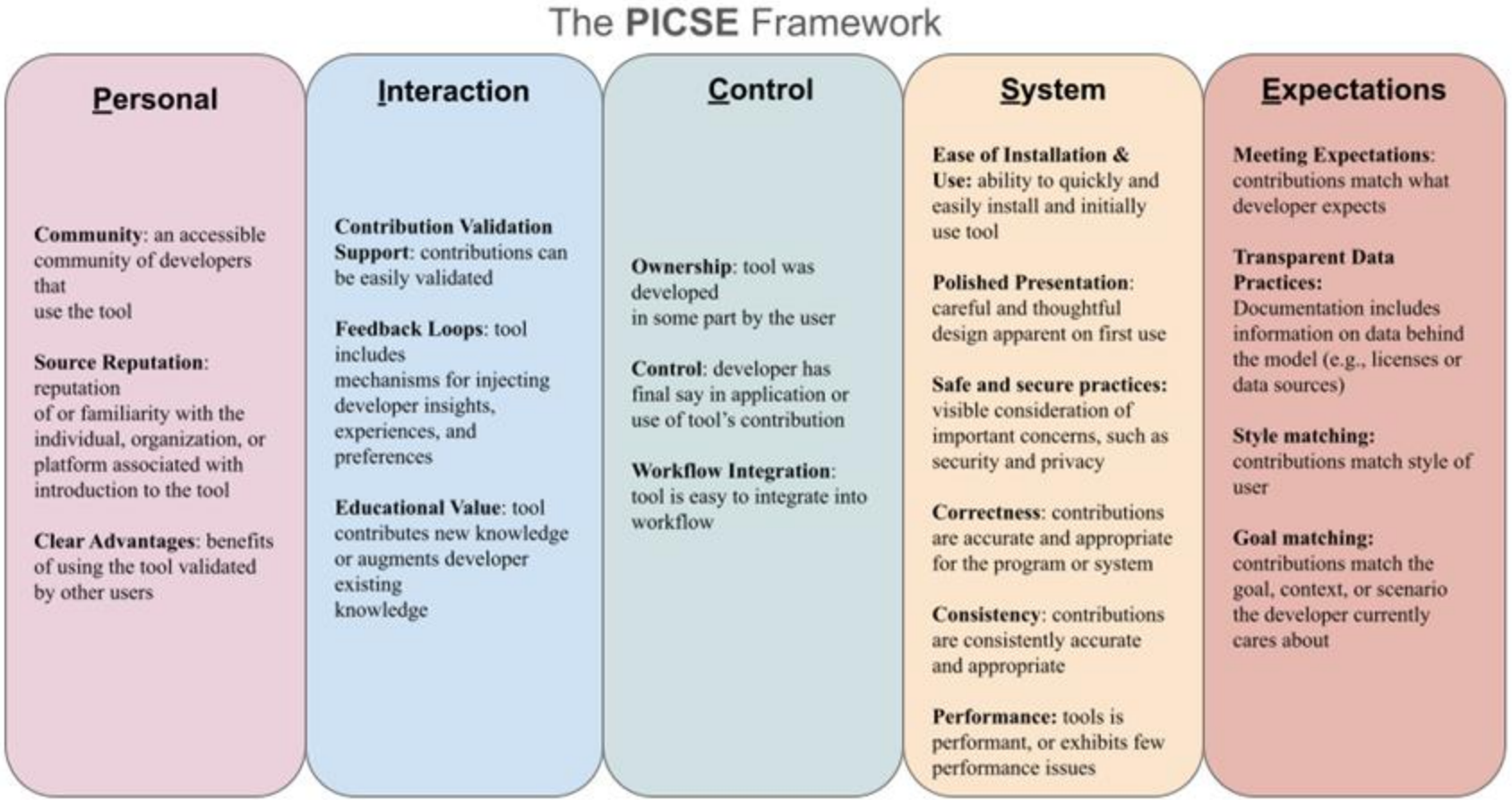
Trust being an ***attitude*** is a psychological construct that is not directly observable & should be:

- captured through *psychometrically validated instruments*
- distinguished from *observable measures such as reliance*

A *foundational design requirement for supporting effective human-AI interactions:*

- Miscalibrated levels of trust can lead developers to:
  - Overlook AI-induced errors and risks in work
  - Eschew its use altogether

*Lee, J. D., & See, K. A. (2004). Trust in automation: Designing for appropriate reliance. Human factors, 46(1), 50-80.*




*Johnson, B., Bird, C., Ford, D., Forsgren, N., & Zimmermann, T. (2023, May). Make your tools sparkle with trust: The PICSE framework for trust in software tools. ICSE-SEIP (pp. 409-419). IEEE.*

# What to prioritize in tool design for trust?

***It is important to establish an understanding of***

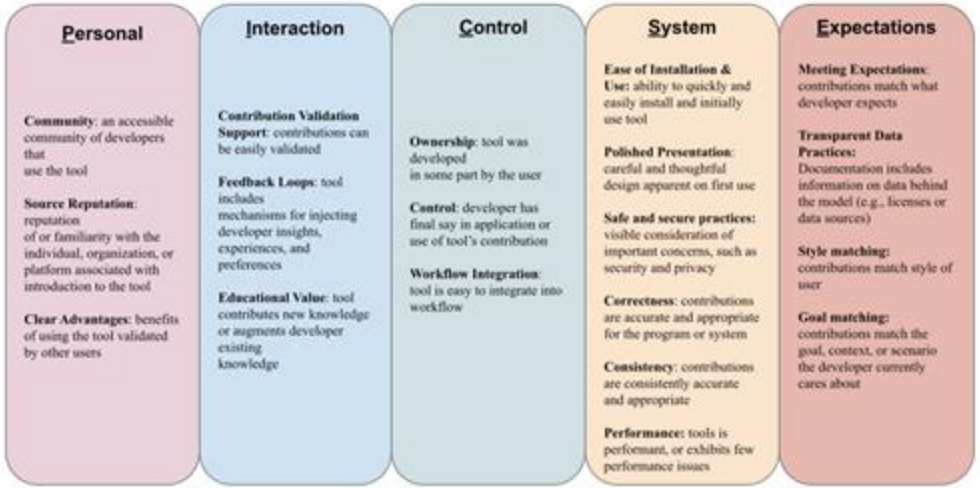
**RQ1: How these *multitude of factors* affect developers' *trust in genAI tools*?**

- A validated instrument for
  - capturing different trust-RELATED factors in human-genAI interaction contexts
    - through a psychometric analysis of the PICSE framework
- The strength & significance of these factors' association with developers' trust in genAI tools

**Survey with software developers (N=238) at GitHub Inc. & Microsoft**

# Psychometric Analysis

***Psychometric quality refers to the objectivity, reliability, and validity of an instrument***




- PICSE was qualitatively developed → Its psychometric quality had not been assessed
- We conducted psychometric analysis of the framework to empirically:
  - refine its factor groupings,
  - which were then evaluated for their association with trust

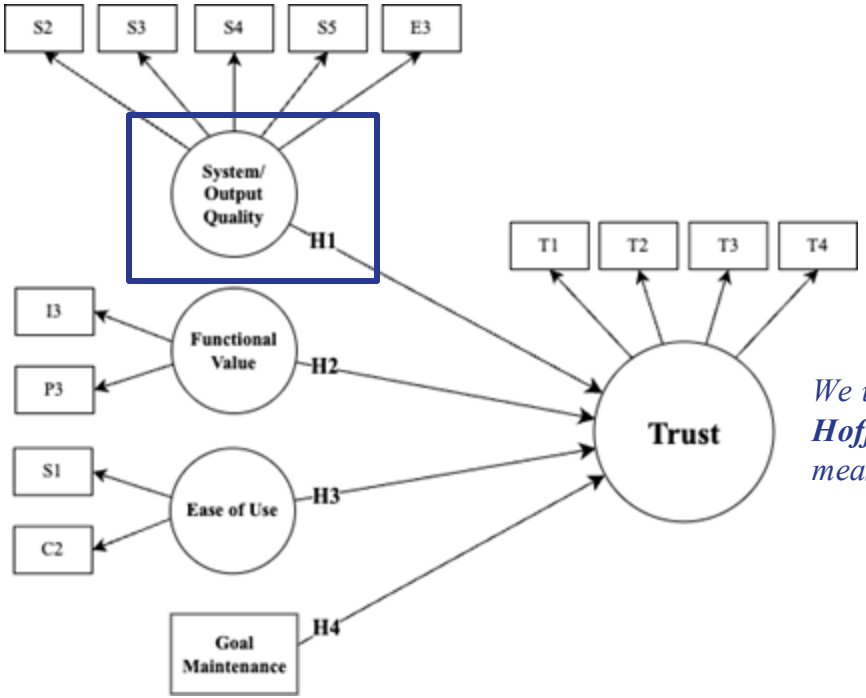
# Validated PICSE Instrument *for capturing trust-RELATED factors in HAI context*

| Construct | ID | Items |
|---|---|---|
| **System/Output Quality** | S2 | Presentation/Interaction design |
| | S3 | Safety/security practices |
| | S4 | Consistent contextual accuracy |
| | S5 | Performance in tasks |
| | E3 | Style matching of contributions |

# Building the structural (theoretical) model

| Construct | ID |
|---|---|
| | I3 |
| **Functional Value** | P3 |
| | S1 |
| **Ease of use** | C2 |
| **Goal Maintenance** | E4 |

*H1: System/Output quality of genAI is positively associated with developers' trust in these tools.*




*We used the validated **Hoffman's TXAI** to measure the **trust construct***

* *Note: TXAI is*

*(a) derived from validated trust scales specifically for HAI interactions,*
*(b) psychometrically validated, and (c) is widely used to capture the trust construct.*

# The exclusionary nature of AI

***often fails to support all users adequately***

Design and Human-Computer Interaction, Language Processing, Machine Learning

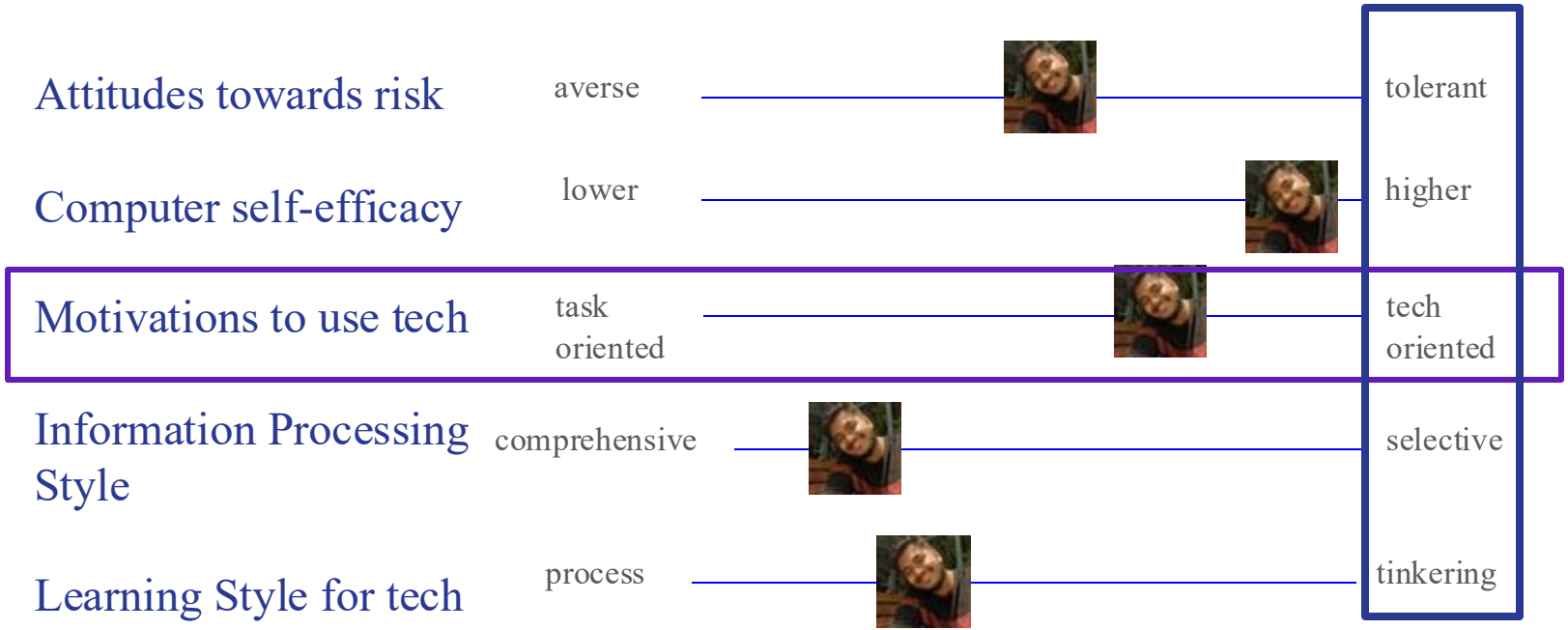
AI-Detectors Biased Against Non-Native English Writers

- Substantial body of work exists in modeling technology adoption,
  - These studies *don't* consider the inclusivity of the software design
  - One such aspect of inclusivity is supporting **cognitive diversity**:
    - Fosters ***divergence in perceptions and interaction styles with technology***
    - No particular style is inherently better or worse
  - When an user's cognitive style is unsupported (or misaligned) by software:
    - Additional "cognitive tax" everytime they use that software
    - Additional barriers to usage and adoption

**RQ2: How are developers' *trust and cognitive styles* associated with their *intentions to use genAI tools*?**

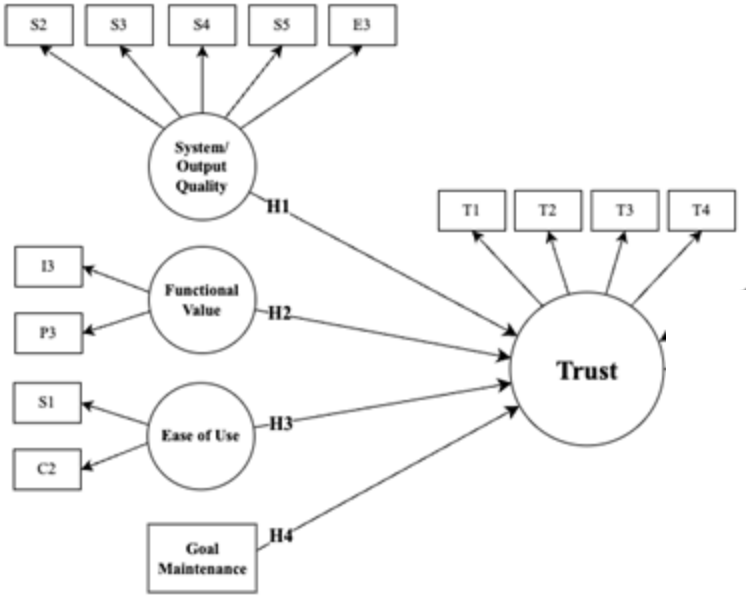
# Cognitive Diversity, *i.e. variations in cognitive styles*

***diverse ways users perceive, process, and interact with information & technology, as well as their approach to problem-solving***



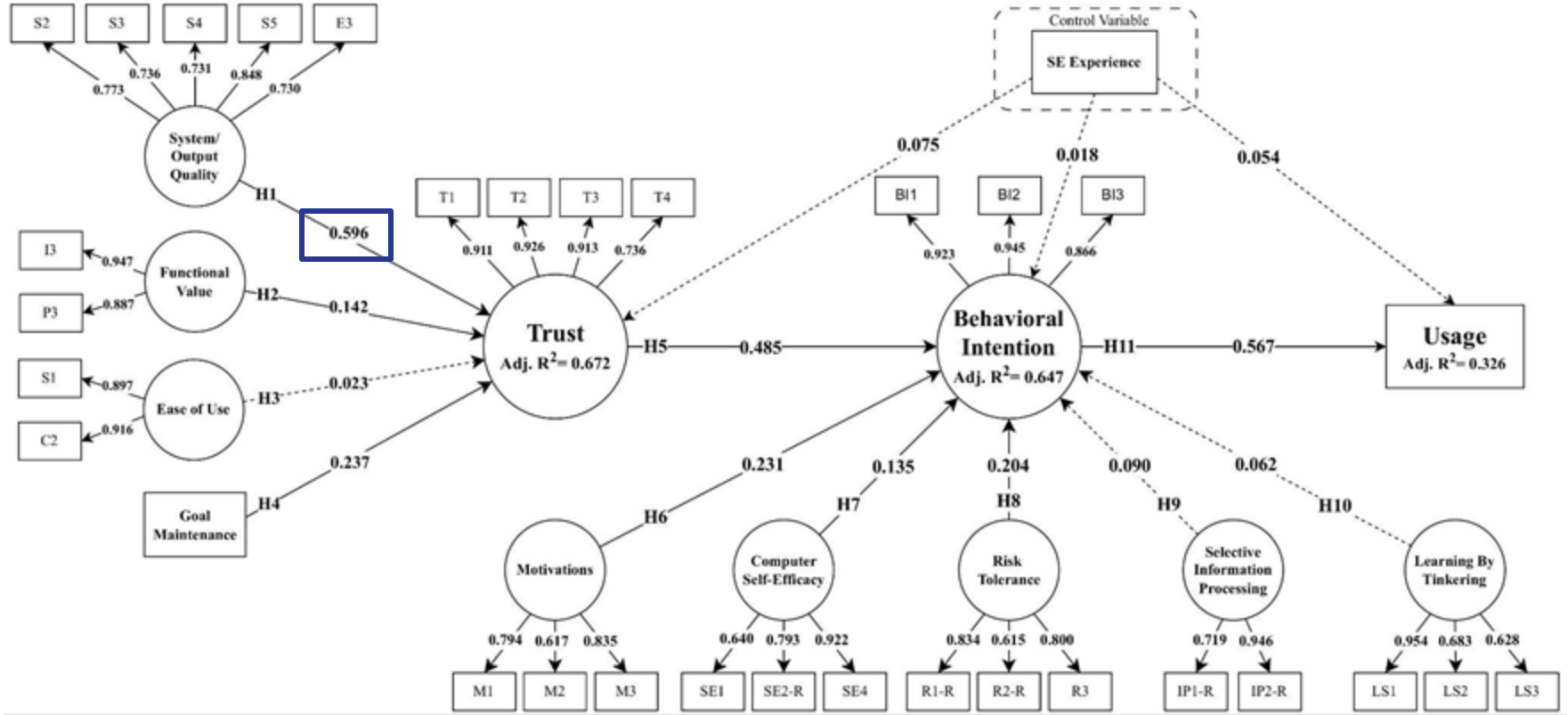

Burnett, M., Stumpf, S., Macbeth, J., Makri, S., Beckwith, L., Kwan, I., ... & Jernigan, W. (2016). GenderMag: A method for evaluating software's gender inclusiveness. *Interacting with computers*, *28*(6), 760-787.

# Building the structural (theoretical) model (contd.)



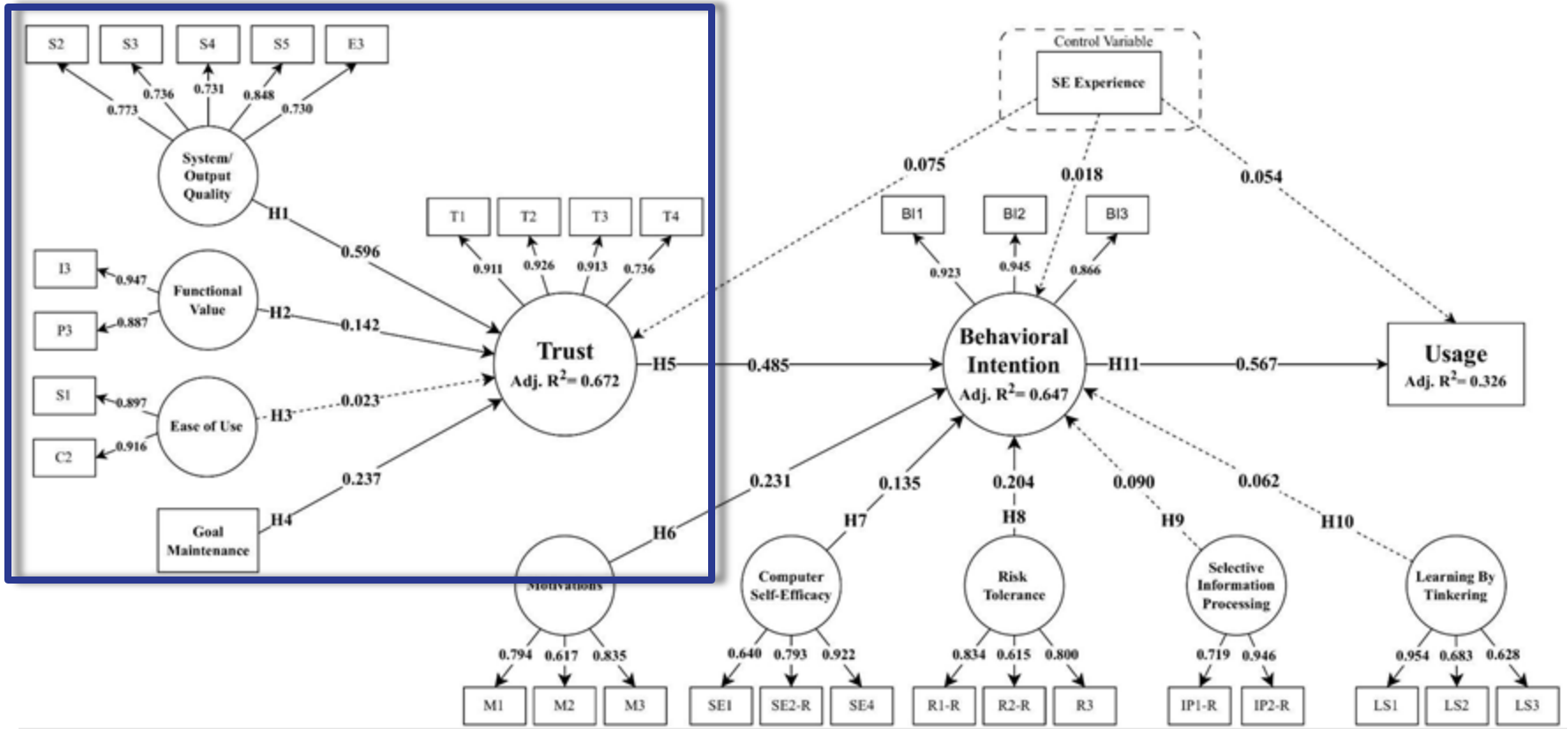

** Note: We used the validated GenderMag facet survey to capture the five cognitive styles*
*We used components of the UTAUT model to capture the behavioral intention and usage constructs*

# Structural Model Evaluation (PLS-SEM)



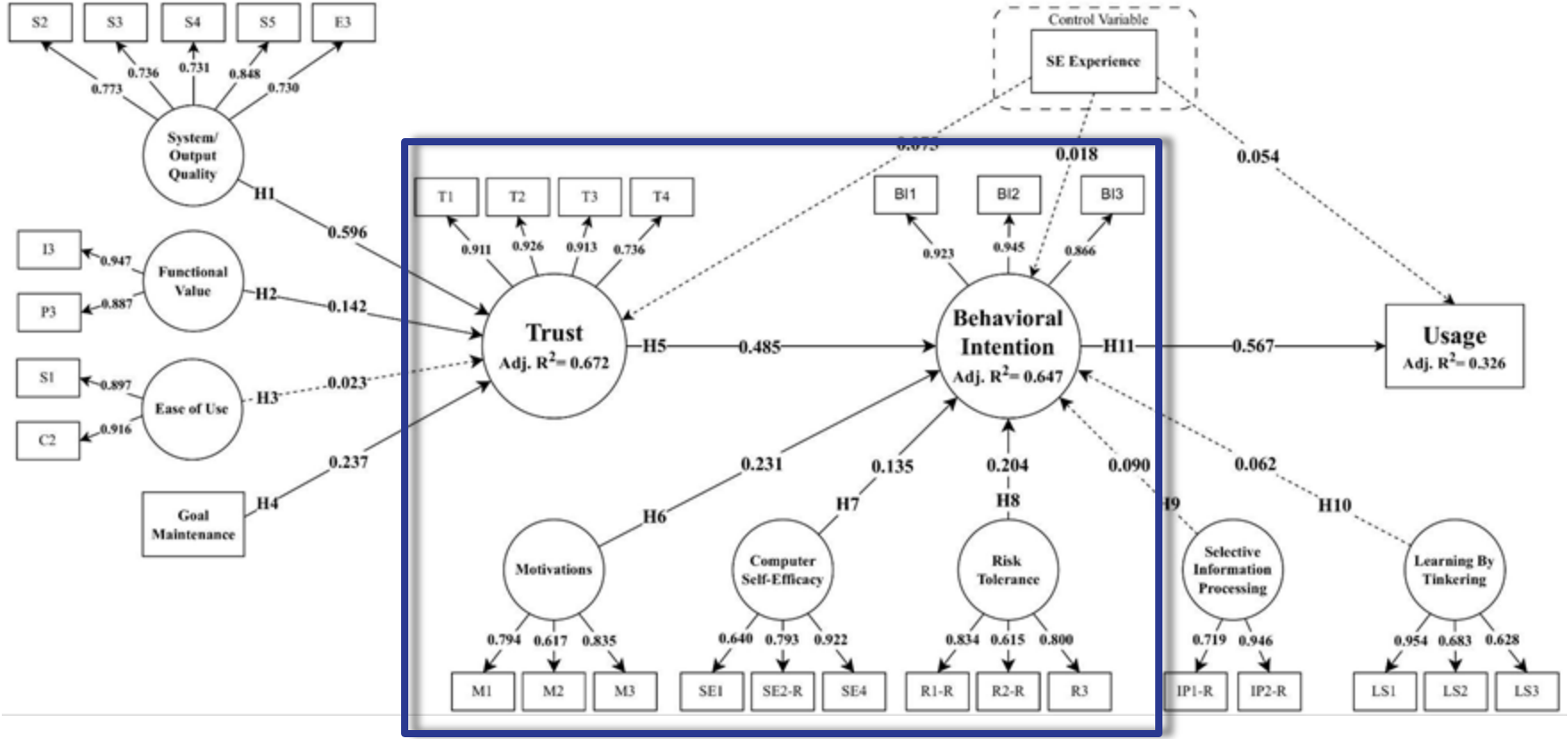

PLS-SEM Model: Solid lines indicate item loadings and path coefficients ($p < 0.05$); dashed lines represent non-significant paths. Reverse-coded items are suffixed with '-R' (e.g., SE2-R). Latent constructs are depicted as circles and adjusted $R^2$ (Adj. $R^2$) values are reported for endogenous constructs.

# Factors associated with trust (RQ1)




PLS-SEM Model: Solid lines indicate item loadings and path coefficients ($p < 0.05$); dashed lines represent non-significant paths. Reverse-coded items are suffixed with '-R' (e.g., SE2-R). Latent constructs are depicted as circles and adjusted R^2 (Adj. R^2) values are reported for endogenous constructs.

# Factors associated with behavioral intentions (RQ2)




PLS-SEM Model: Solid lines indicate item loadings and path coefficients ($p < 0.05$); dashed lines represent non-significant paths. Reverse-coded items are suffixed with '-R' (e.g., SE2-R). Latent constructs are depicted as circles and adjusted R^2 (Adj. R^2) values are reported for endogenous constructs.

# Takeaways: Prioritizing drivers of trust and adoption

Use cognitive factors to guide design improvements

Prioritize designing for goal maintenance, contextual transparency, & inclusive HAI-UX

Leverage the model &/or the validated scale in your context to improve understanding of AI adoption dynamics in SE

Important to design these tools to not only assist with tasks but also meaningfully support the people who use it.

# Thank You!

*Questions?*

choudhru@oregonstate.edu

Check out our paper!

Oregon State University