eurostat

International Labour Organization

Global Conference on Measuring New Forms of Employment

Brussels, 4 - 5 July 2024

#MeasuringEmployment

GENERATIVE AI AND THE FUTURE OF EMPLOYMENT: A COLLABORATIVE APPROACH

Session 6

Professor Aleksandra Przegalinska
Kozminski University, Harvard, CampusAI

THE FUTURE OF WORK WITH COLLABORATIVE AI

WHY COLLABORATIVE AI? (IT'S MORE THAN HUMAN-CETRIC)

EVOLUTION

LARGE LANGUAGE MODELS + DIFFUSION MODELS

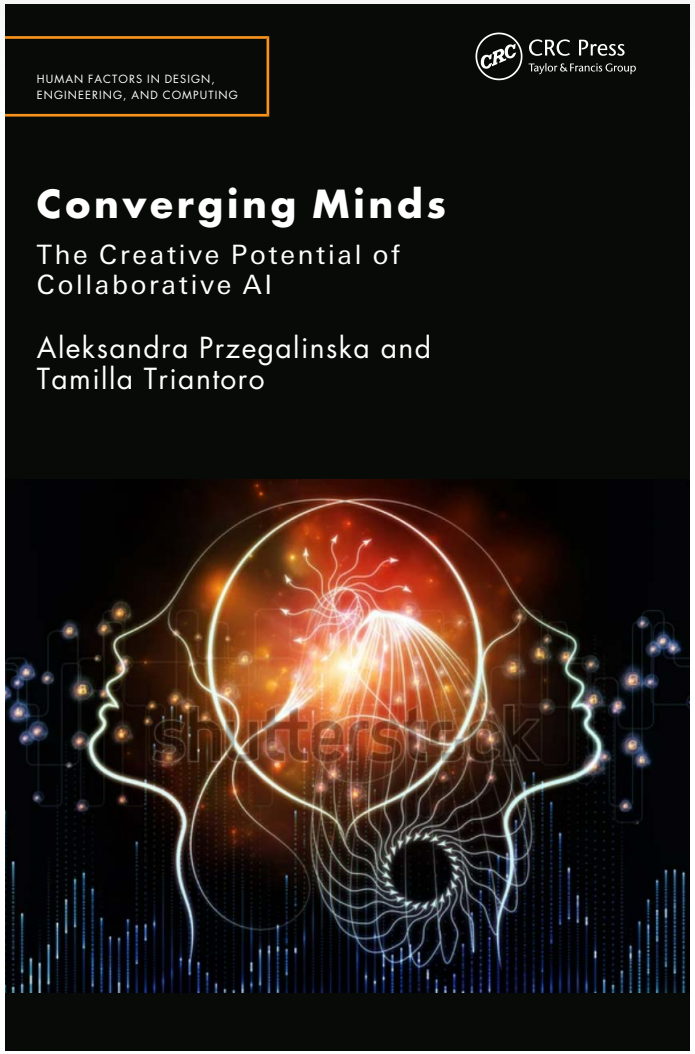
CONVERSATIONAL MODELS

MULTIMODAL MODELS

LARGE ACTION MODELS

OMNIMODELS

New book written with Claude, Bard and GPT



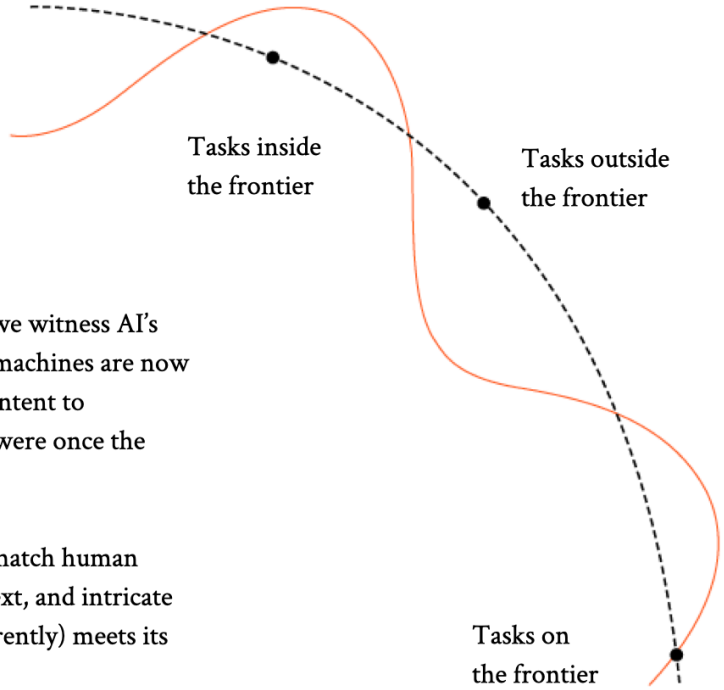
ASSISTANTS (NOT AGENTS):

- CRITIC
- STYLE AMPLIFIER
- IDEA GENERATOR
- FIELD EXPERT

It greatly elevated our academic output!

WE ALREADY HAVE EVIDENCE CONCERNING AI AND PRODUCTIVITY

THE JAGGED FRONTIER



The jagged frontier works as follows: on one end of the spectrum, we witness AI's remarkable prowess — tasks that once seemed insurmountable for machines are now executed with precision and efficiency (from generating creative content to predicting complex patterns), with AI showcasing capabilities that were once the exclusive domain of human cognition.

Yet, on the flip side, there are tasks where AI falters, struggling to match human intuition and adaptability. These are areas marked by nuance, context, and intricate decision-making — realms where the binary logic of machines (currently) meets its match.

7

OUR FORMER STUDY

COBOTS IN THE KNOWLEDGE WORK:

40% - job effectiveness increase
60% - job satisfaction increase

https://www.sciencedirect.com/science/article/pii/S0148296320307 92X

THEORIES:

- ***Resource-Based View*** *- dependence of an organization's competitive edge on the unique resources it owns or controls*

- ***Task Technology Fit****- provides a model that details how the characteristics of a task, paired with the attributes of the technology used to perform it, can influence individual performance.*

HYPOTHESES

- *H1: Firms that possess advanced generative AI resources and capabilities will have a competitive advantage over firms that do not.*
- *H2: Firms with employees more experienced in AI will have a competitive advantage over firms that do not.*
- *H3a: Firms utilizing generative AI for automation tasks will perform these tasks more effectively than firms that do not.*
- *H3b: Firms employing generative AI for decision support tasks will perform these tasks more effectively than firms that do not.*
- *H3c: Firms that integrate generative AI in their creative tasks will perform these tasks more effectively than firms that do not.*
- *H3d: Firms that leverage generative AI for innovation tasks will perform these tasks more effectively than firms that do not.*

Table II. Task classification based on complexity and creativity

Task	Simple	Complex
Routine	Automation task: Persona Ideation	Decision Support Task: Competitive Analysis
Creative	Creation task: Text-based Ad	Innovation task: Product Naming

Participants n=94

Participants from Executive MBA groups were experienced and had seniority in management related tasks, including marketing.

The participants completed the tasks within a set time frame (45 mins).
On average participants completed study within 19.5 min (SD = 10.8 min).

INDEPENDENT JUDGES QUALITY ASSESMENT

The quality of the output for each task was evaluated by a panel of independent judges rating participants' responses for the tasks (scale from 1 to 5).

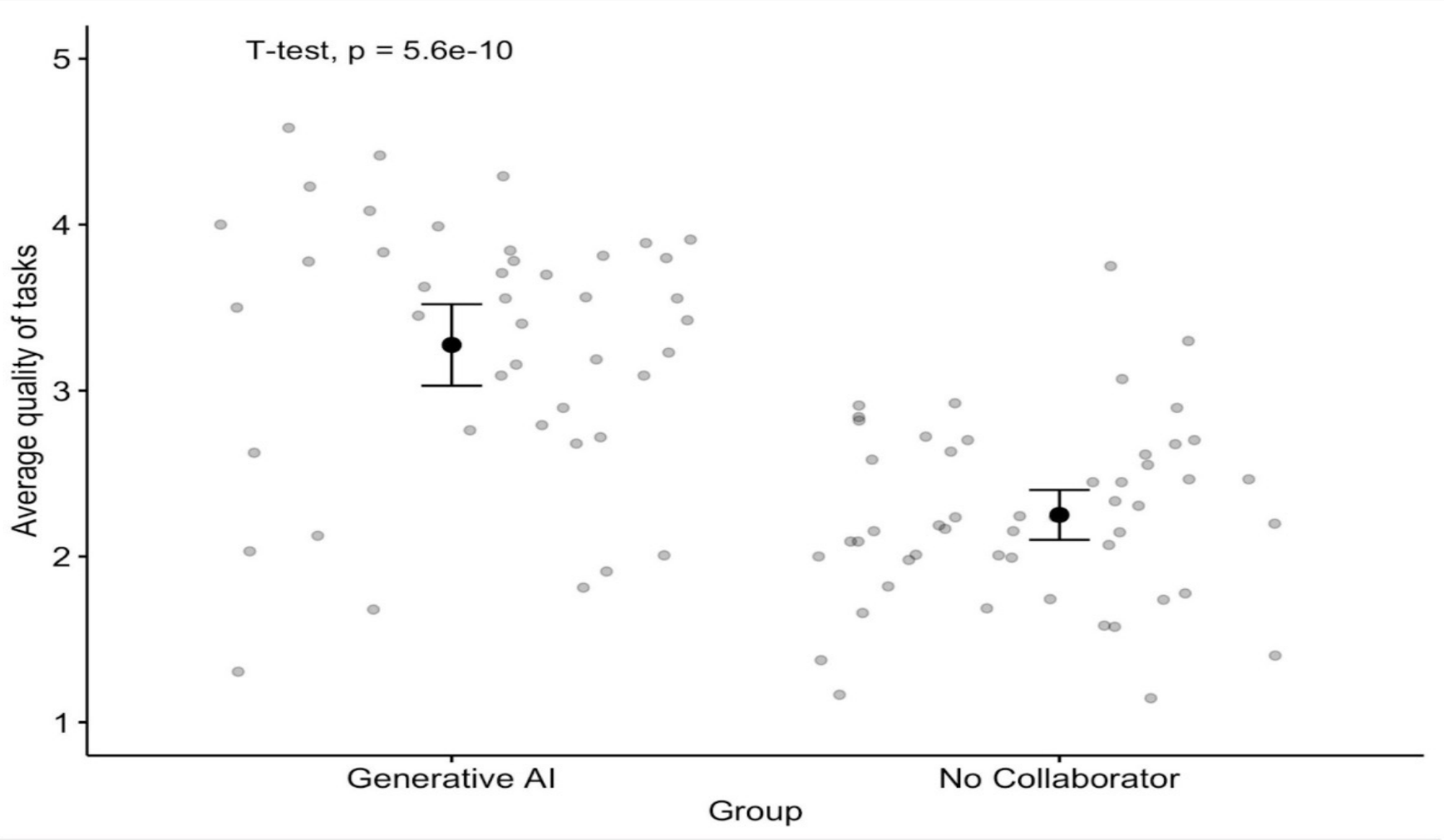
Cronbach's alpha was calculated to measure the **reliability of the created scales.** Cronbach's alpha for the quality of product name scale was 0.84, competitive analysis 0.98, text-based ad 0.97 and persona 0.98

The methodology also included qualitative review of experiment output by the judges. The results of the analyses revealed an average inter-judge reliability of 0.464 (Krippendorf Alpha) which indicates a moderate level of agreement among the judges.

HYPOTHESES TESTING

Hypothesis 1:

*Welch Two Sample t-test to examine the difference **in the quality of task output** between participants in the group interacting with Generative AI (M = 3.27, SD = 0.8, n = 43) and the group with No Collaborator (M = 2.25, SD = 0.54, n = 51)*


T-test, p = 5.6e-10
Average quality of tasks
5
4
3
2
1
Generative AI
No Collaborator
Group

Hypothesis 2:

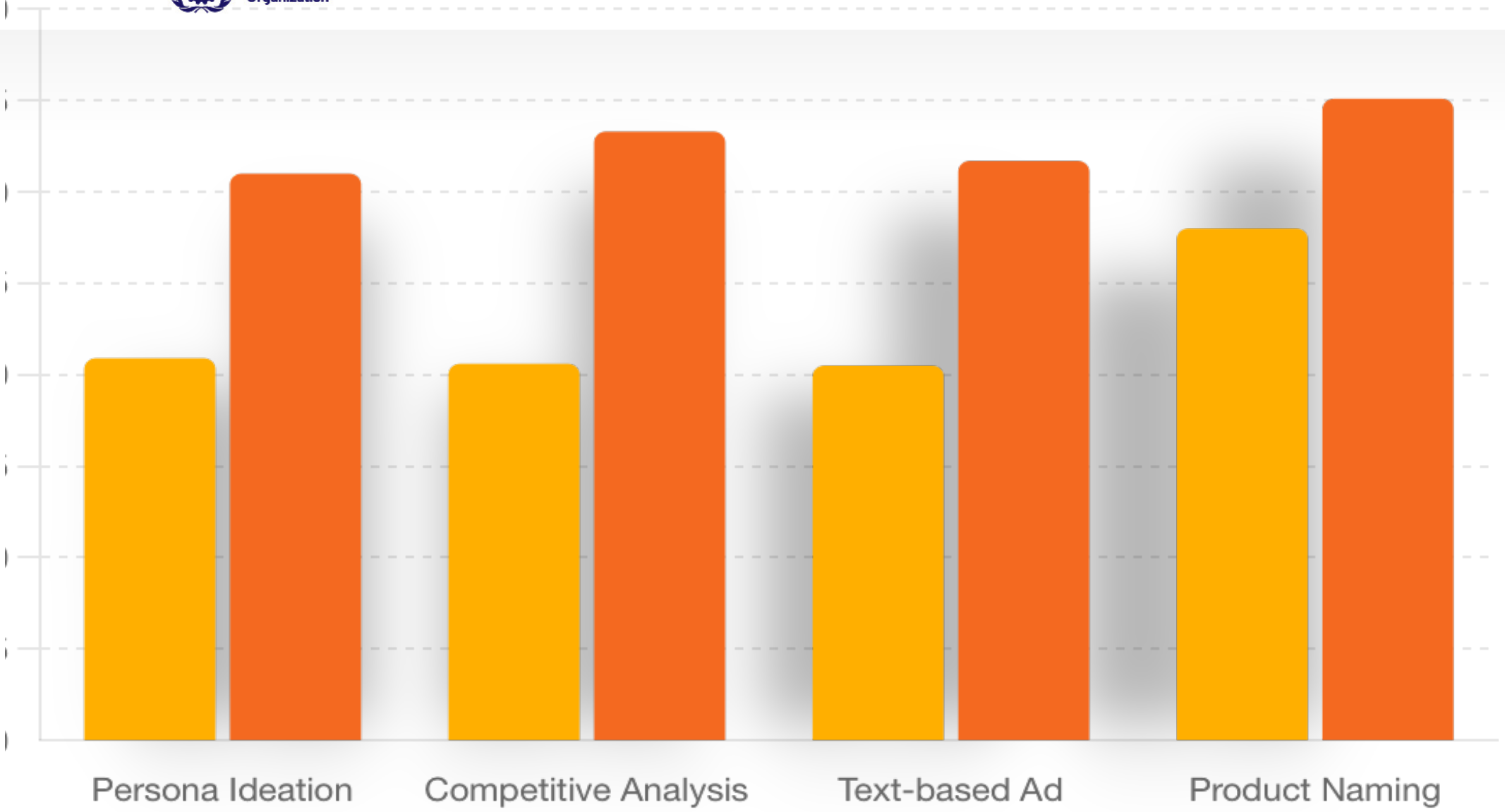
Welch one-way ANOVA to investigate the ***differences in the quality of the tasks between the groups with no familiarity, moderate familiarity, significant familiarity with Generative AI.***

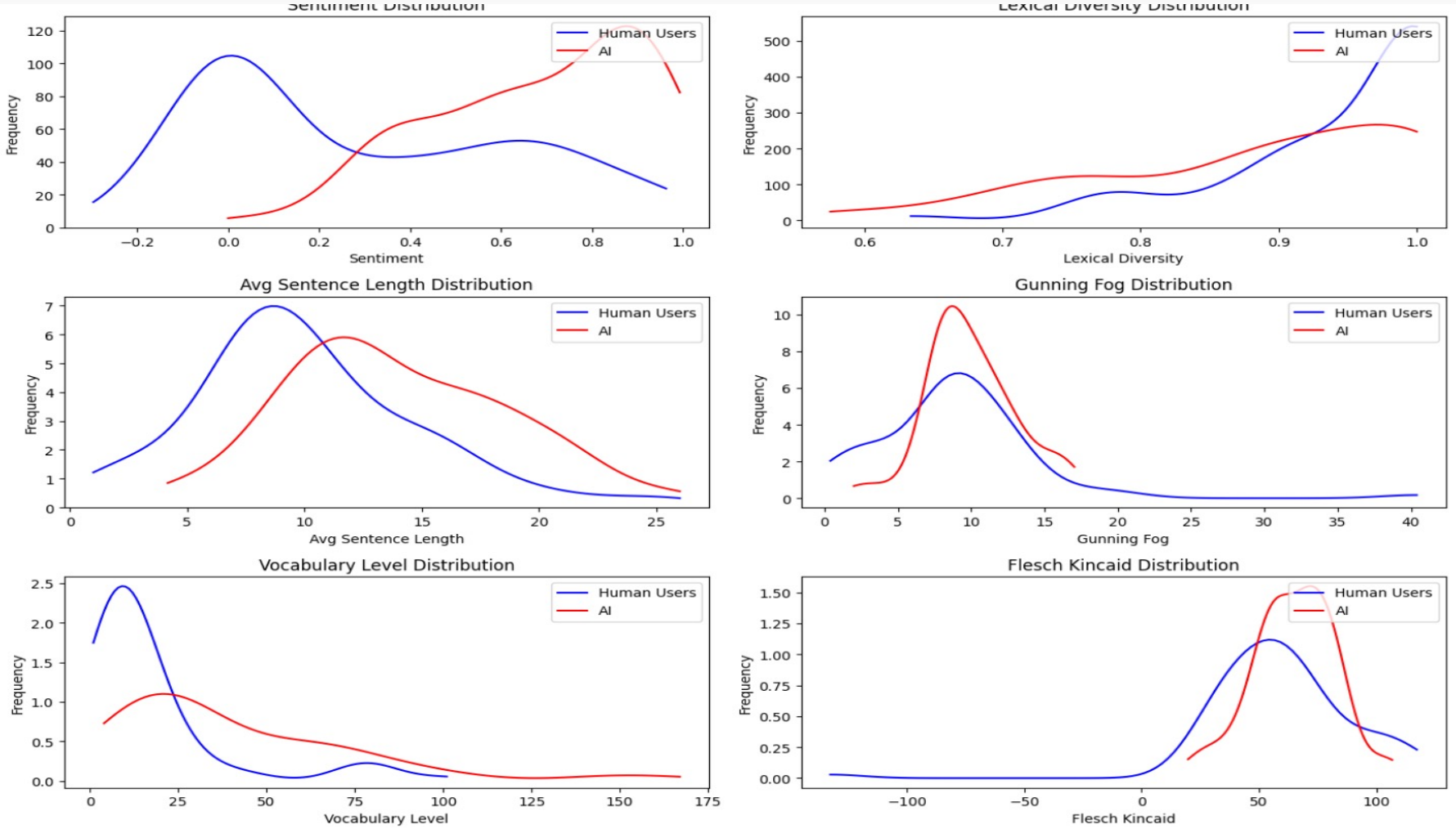
Table III. Descriptive Statistics for Quality of Tasks Completed by Participants with No familiarity Moderate familiarity, Significant familiarity with Generative AI.

Familiarity with Generative AI		Persona Ideation				Competitive Analysis			Text-based Ad			Product Naming		
	N	*M*	*SD*	*F*	*M*	*SD*	*F*	*M*	*SD*	*F*	*M*	*SD*	*F*
No familiarity	12	1.95	0.75	5.18**	1.99	0.84	4.77*	2.10	0.79	3.77*	2.89	0.56	0.44
Moderate familiarity	64	2.54	1.01		2.65	1.08		2.53	0.96		3.14	0.79	
Significant familiarity	13	3.23	1.13		3.31	1.18		3.12	0.96		3.31	0.63	

*The effect size was medium for **product name** and large for **competitive analysis** and **persona ideation tasks.***

*The medium effect size indicates that there is a moderate, but still noticeable, relationship between the familiarity with Generative AI technology and the quality of the **text-based ad.***


Persona Ideation
Competitive Analysis
Text-based Ad
Product Naming


Sentiment Distribution
Human Users
AI
Frequency
Sentiment
Lexical Diversity Distribution
Lexical Diversity
Avg Sentence Length Distribution
Avg Sentence Length
Gunning Fog Distribution
Gunning Fog
Vocabulary Level Distribution
Vocabulary Level
Flesch Kincaid Distribution
Flesch Kincaid

UNDERSTANDING HUMAN-AI PATTERNS

AI VS. PERSONALITY

Methodology

- **Participants:** Varied in age, gender, and employment status.
- **Tasks:** Marketing campaign tasks similar to Study 1
- **Metrics Evaluated:** Task quality, engagement duration, and future AI usage intentions.

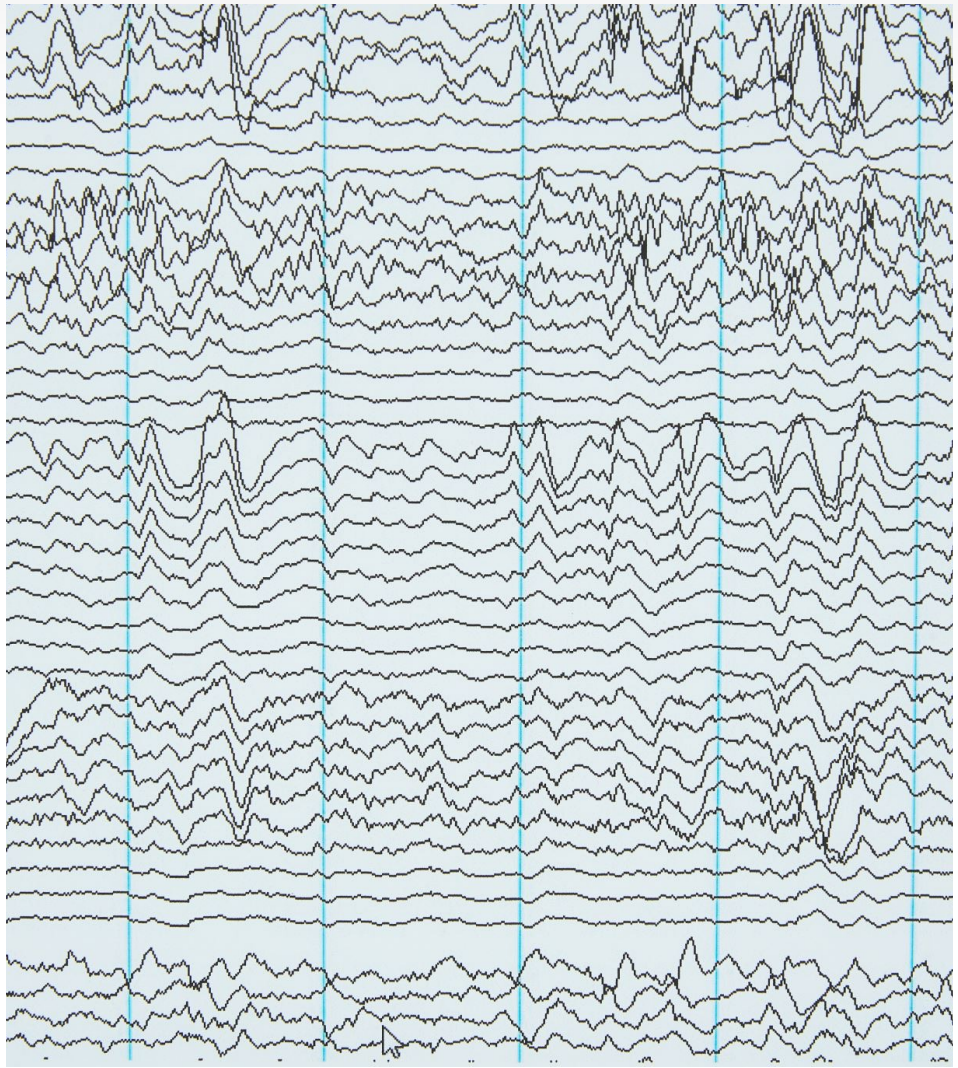
Methodology

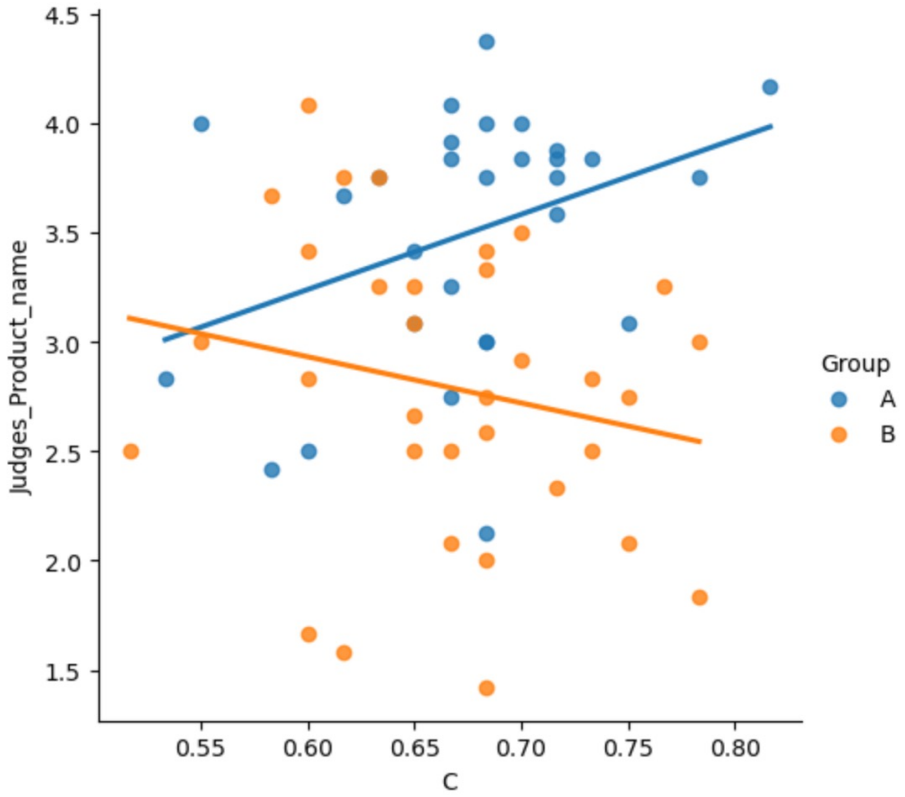
- We used *lsr* library and eta squared function for the calculation of effect size (2) and partial effect size (p2).
- Effect sizes were interpreted based on established guidelines:
- p2 around 0.01 - small,
- p2 around 0.06 - medium,
- p2 around 0.14 and higher - large
- (Cohen 2013; Miles and Shevlin 2001)

- Following the ANOVA, post hoc analyses including simple slopes and pairwise comparisons were performed using *interactions* and *emmeans* libraries to investigate specific group differences.

- In cases when interaction with the group was insignificant, linear regression analysis using ordinary least squares method (OLS) was conducted to predict performance based on traits.

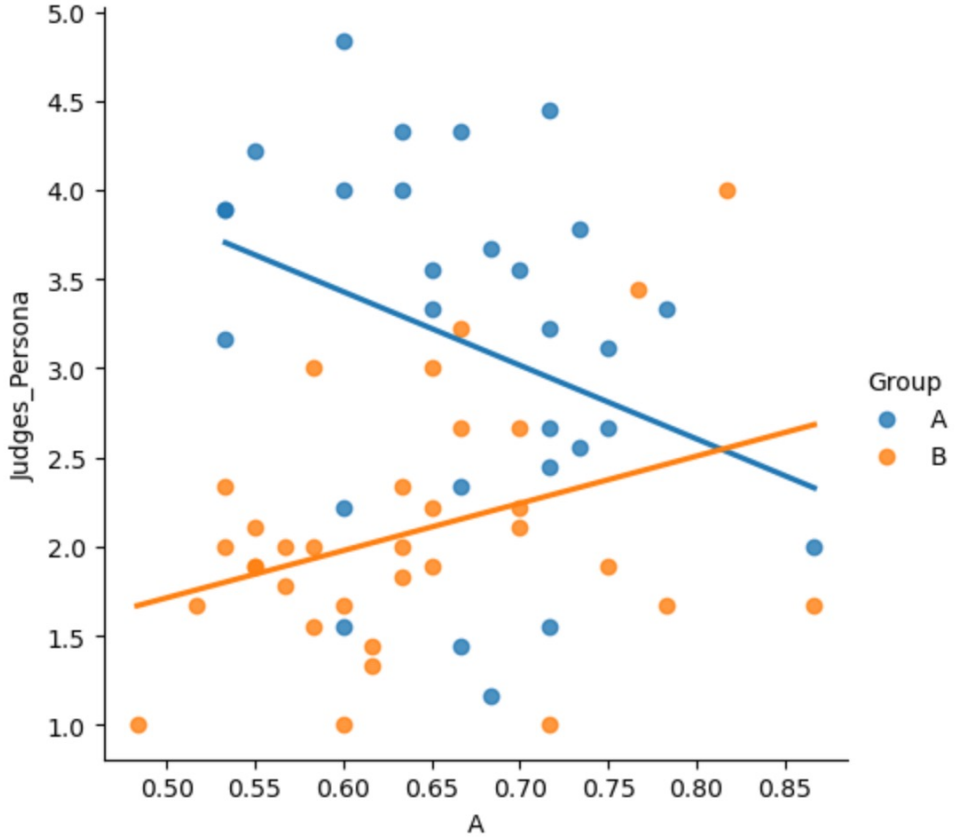
- Additionally such libraries as *dplyr* and *ggplot2* were used for data exploration and visualization.





BIG FIVE RESULTS:

"People who are more conscientious and collaborate with a chatbot achieved better quality in persona creation, while less conscientious people who do not collaborate with the bot had better results for product name quality."



BIG FIVE RESULTS:

"Less agreeable individuals collaborating with a chatbot achieved better quality in all tasks"

Implications for AI Design

- Incorporate psychological principles into AI design for better user alignment.
- Tailored AI solutions can enhance user satisfaction and productivity.

www.campus.ai

eurostat

International Labour Organization

Thank you!

Global Conference on Measuring New Forms of Employment

Brussels, 4 - 5 July 2024