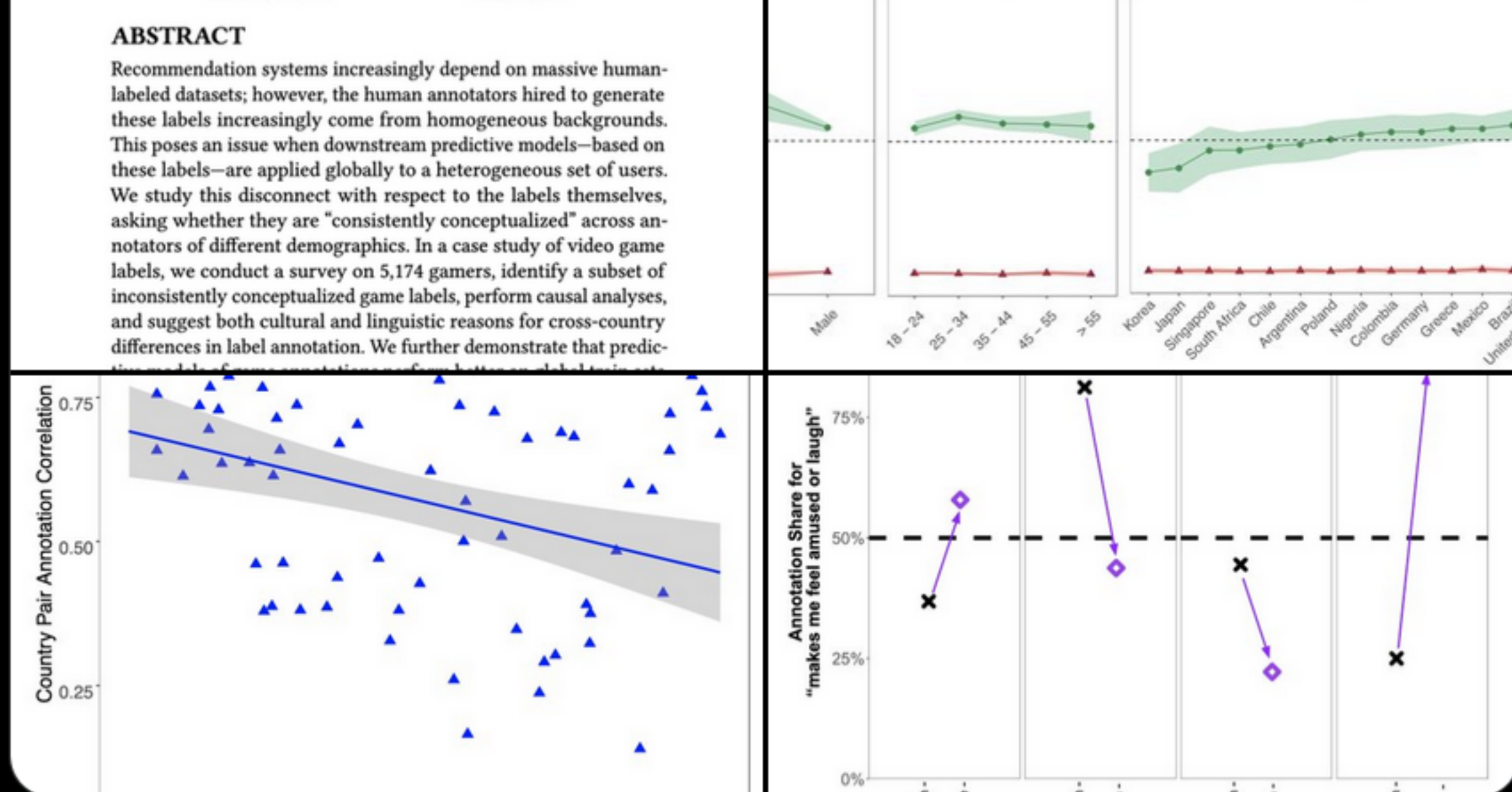




Allison Koenecke @allisonkoe

Had a great time collabing w/ @MSFTResearch & @Xbox on @rockpang6's internship w/ @KatieStonePerez, resulting in a new @FacCTConference paper! "Auditing Cross-Cultural Consistency of Human-Annotated Labels for Recommendation Systems"! Link: dl.acm.org/doi/10.1145/35...; 📄 📌 (1/9)



Microsoft Research New England and 3 others

3:05 PM · Jun 14, 2023 · 11.3K Views

2 replies, 13 retweets, 66 likes, 12 bookmarks

Reply

Allison Koenecke @allisonkoe · Jun 14, 2023
Our research question is: how do we better understand cross-cultural differences in label annotations, specifically how humans label different video games? 🎮 Labels such as "action"/"replayable"/"anime"/"difficult"/"cozy" are useful for both organizing and discovering games (2/9)

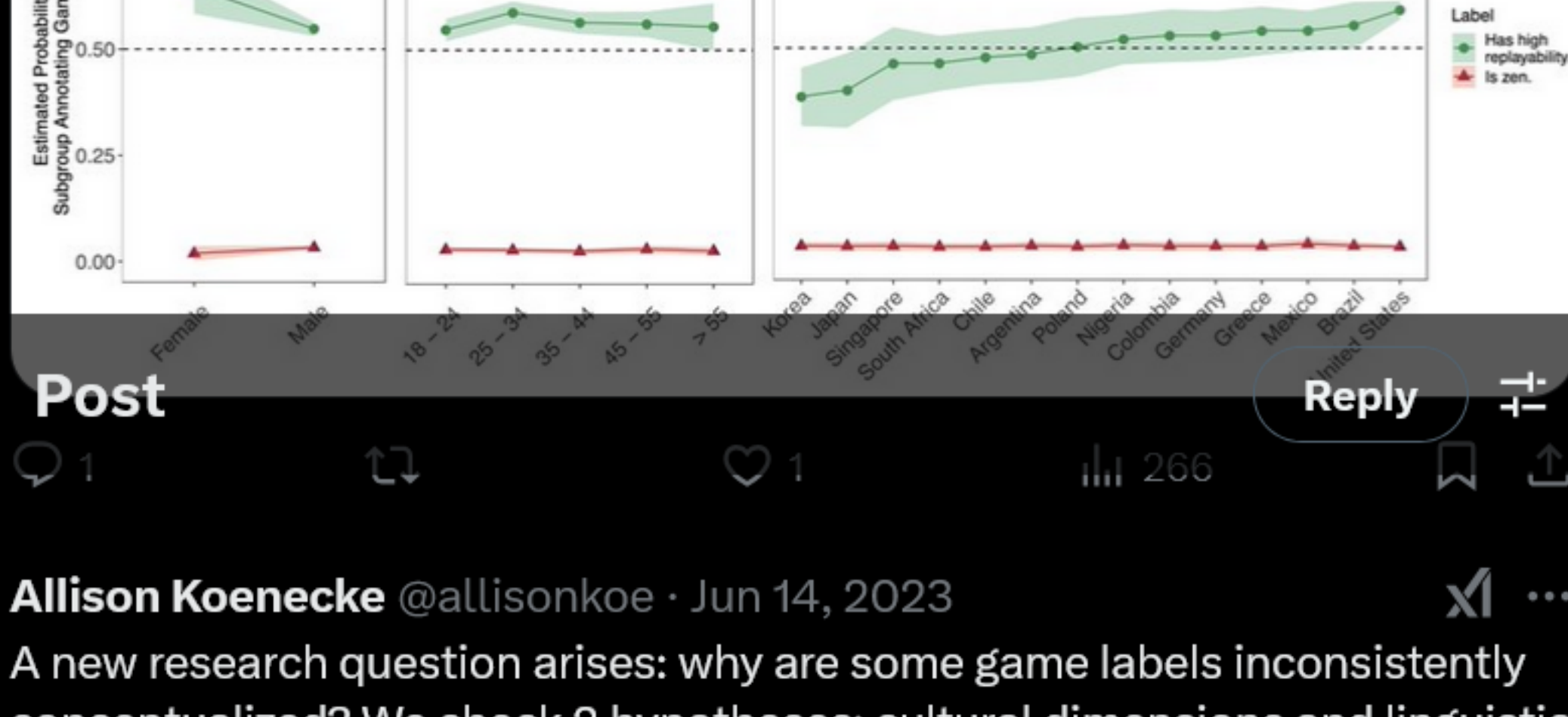
Allison Koenecke @allisonkoe · Jun 14, 2023
Our partner on this project is Microsoft's Gaming for Everyone team, whose mission is to create more diverse game experiences on Xbox and appeal to a broader set of customers. They want to ensure gaming discovery is inclusive and understand potential biases in their system (3/9)



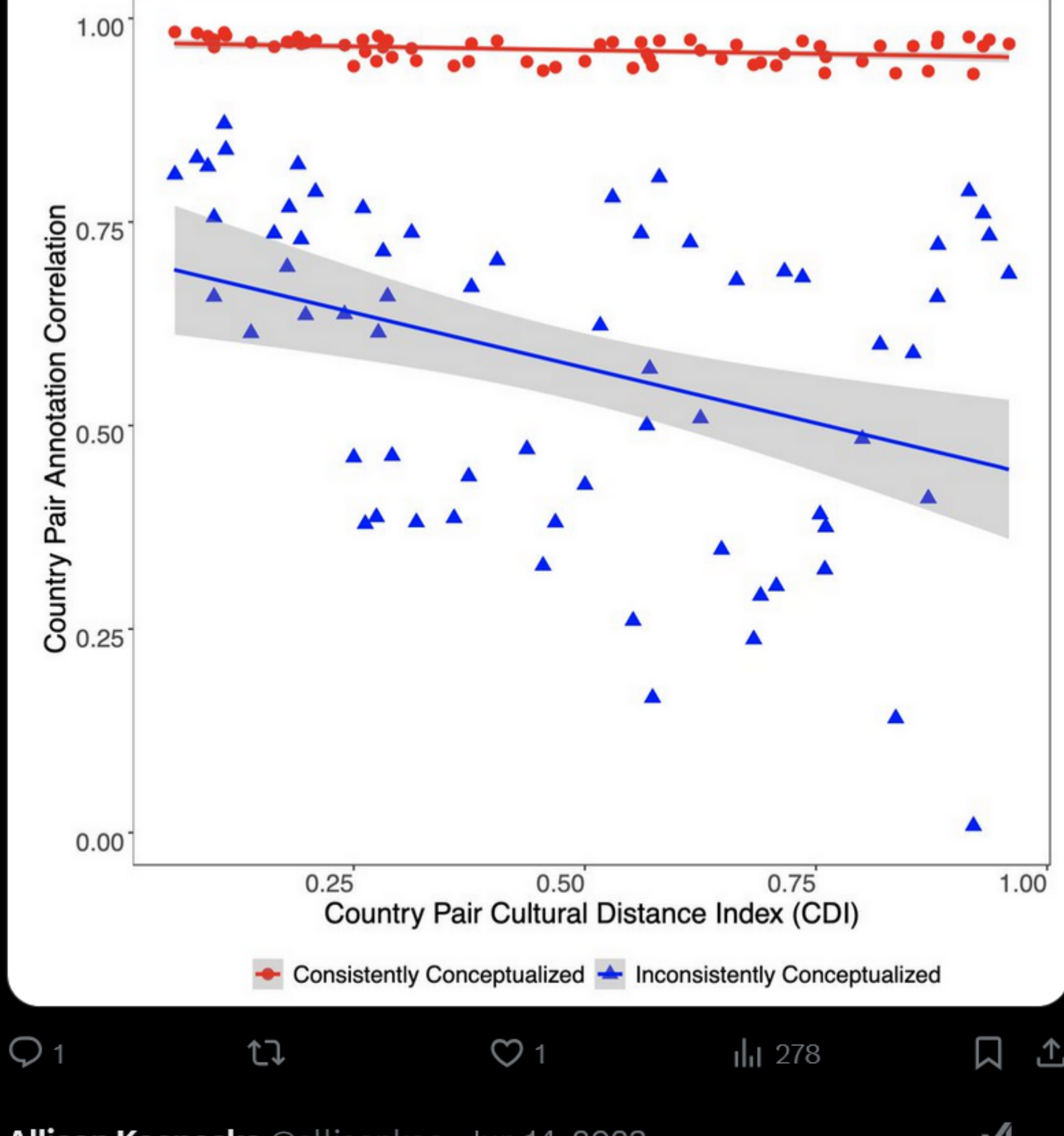
1 reply, 1 retweet, 2 likes, 413 views

Allison Koenecke @allisonkoe · Jun 14, 2023
To audit biases in gaming discovery, we ran a large-scale survey of gamers (including @XboxAmbassadors) in 16 countries and 9 languages to understand how different gamers label the same games. We asked respondents which of 28 game labels would apply to 11 video games (4/9)

Allison Koenecke @allisonkoe · Jun 14, 2023
We classify game-label pairs as either consistently conceptualized across cultures (eg. everyone agrees that a certain action game will not be labeled "zen") or inconsistently conceptualized (eg. majority of US thinks this game is "highly replayable", not true for Koreans) (5/9)

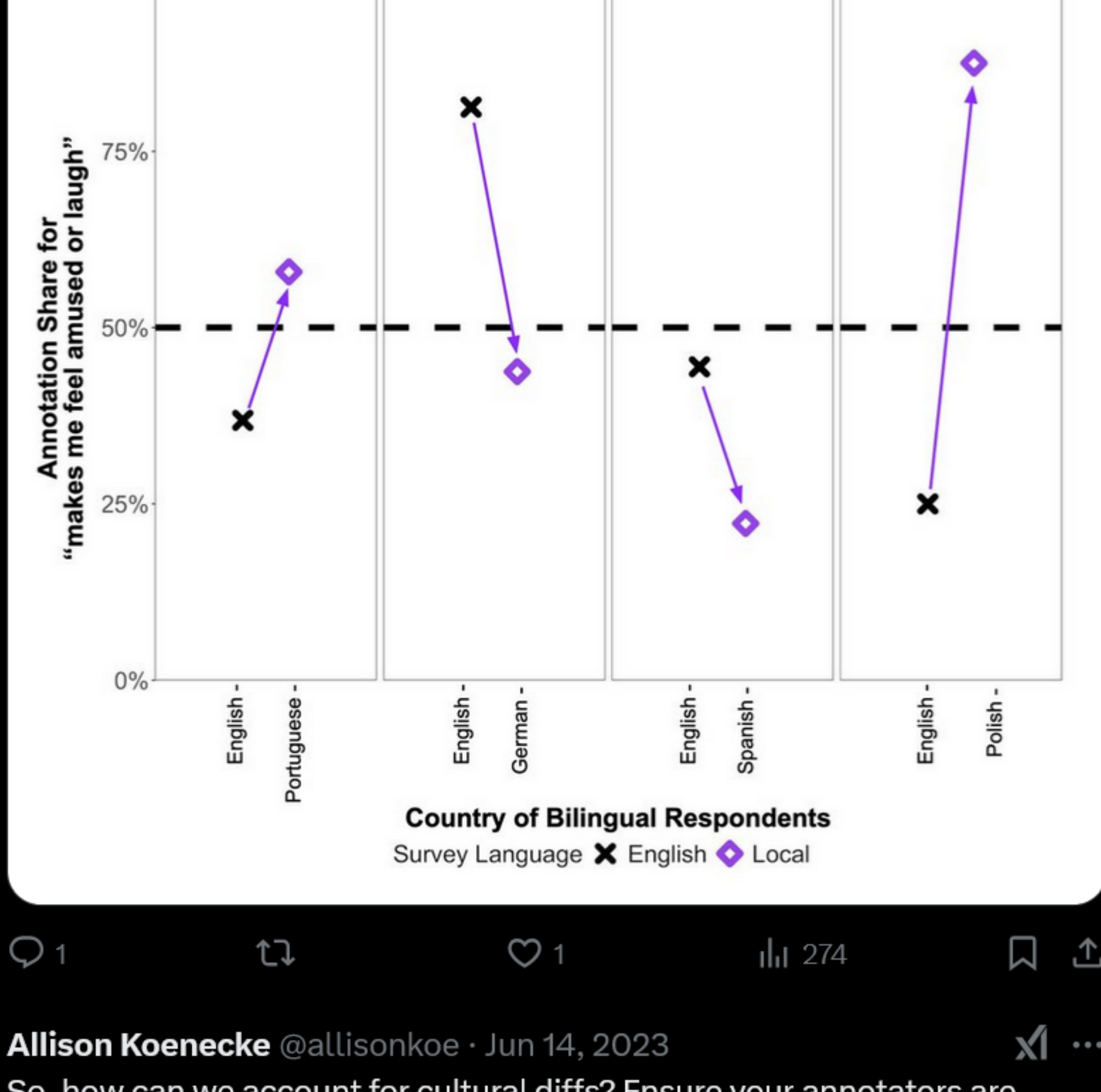


Allison Koenecke @allisonkoe · Jun 14, 2023
A new research question arises: why are some game labels inconsistently conceptualized? We check 2 hypotheses: cultural dimensions and linguistic differences. First, countries culturally further away (based on Hofstede dimensions) do tend to label games more differently (6/9)



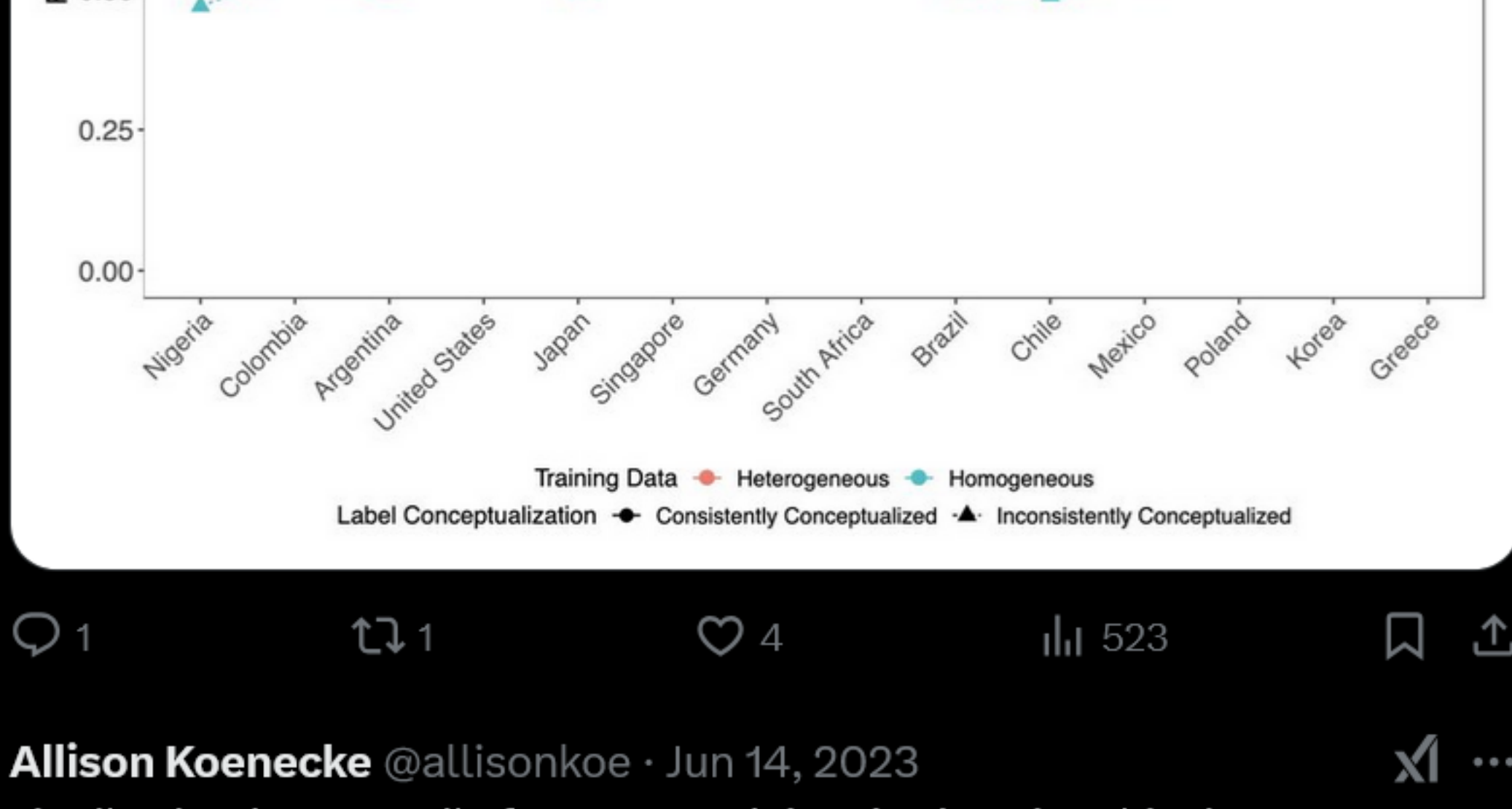
1 reply, 1 retweet, 1 like, 278 views

Allison Koenecke @allisonkoe · Jun 14, 2023
Second, translation plays a key role in conceptualizing label meaning! We performed a causal matching analysis on bilingual speakers and found that the same label in different languages results in different likelihoods of being associated with a game (7/9)



1 reply, 1 retweet, 1 like, 274 views

Allison Koenecke @allisonkoe · Jun 14, 2023
So, how can we account for cultural diffs? Ensure your annotators are diverse (➡️ predictive models perform better across the board)! More in the paper: a step-by-step framework for auditing your own data labels & the need for more qualitative evaluations of cultural diffs (8/9)



2 replies, 1 retweet, 4 likes, 523 views

Allison Koenecke @allisonkoe · Jun 14, 2023
Finally, thank you to all of our research inspirations in ethical crowdsourcing, & the broader gaming/fairness/HCI/recsys community! @nancybaym @marylgray @ssuri @turkopticon @alexhanna @MilagrosMiceli @blahtino (9/9)

Allison Koenecke @allisonkoe · Jun 15, 2023
+ extra thanks to @o_saja for helping get this paper off the ground!!