

## Do you perceive certain Reddit usernames as belonging to a male or female?

- n = 1335 (wow!)
- Open for a little over 24 hours

My investigation into this question consisted of 2 parts:

- Initial survey asking respondents to provide their Reddit username and gender (male/female only)
  - Ended up having **693** legitimate responses (e.g. not duplicates, provided username *and* gender, didn't give an "other" gender). Gender breakdown was **40% M** and **60% F**.
- Follow-up survey with **50** randomly selected usernames from the above survey, where respondents were asked whether they thought each username belonged to a male or female. *This is what I'm analyzing in this document!*
  - The gender breakdown for usernames here was **38% M** and **62% F**, meaning the "population" described above yielded a representative sample. However, that doesn't mean this ~40/60 ratio holds true for all r/SampleSize users. It's possible my survey appealed more to female users than the makeup of the sub as a whole.

## How right or wrong were respondents?

### Summary

- 17 male** usernames were guessed **correctly**. (89.5%)
- 14 female** usernames were guessed **correctly**. (45%)
- 2 male** usernames were guessed **incorrectly**. (10.5%)
- 17 female** usernames were guessed **incorrectly**. (55%)

### By Username

The text color indicates the actual gender of that username.

The M/F next to the number indicates % that guessed this gender.

### Right

n = 31

- **Strongly predicted correct gender (80.0% of respondents and above)**
  - **ThalassaTea** (96.0F), **OJbutNotOJSimpson** (95.2M), **Tapemaster21** (94.2M), **Gamecubed-** (92.4M), **ThatSucc** (89.7M), **gamerz1818** (89.6M), **\_maybe** (88.7F), **Lansklo** (85.2M), **GreenspaceCatDragon** (83.1F), **Kupfernitrat** (82.9M), **Thicc\_hung\_nBisexual** (80.8M)
- **Moderately predicted correct gender (60.0-79.9% of respondents)**
  - **tmrika** (79.0F), **whinydog** (77.5M), **zwiiz2** (76.0M), **kaz3ro** (70.5M), **TheCuriousProgram** (69.5M), **Chicken-Cat** (67.0F), **Curiositygotthecat99** (67.0F), **Texas\_Toast** (66.7M), **Zensandwich** (64.6F), **Mindssofaraway** (64.4F), **Midwestthistle** (62.0F), **Booyahbitches** (60.1F)
- **Slightly predicted correct gender (51.1-59.9% of respondents)**
  - **InterrrogatingTheCat** (57.5F), **Doninari** (55.7M), **grayturtle31** (55.7M), **bendtowardsthesun** (53.6F), **combatwombat02** (53.3M), **Sweatytiits** (52.8F), **volthicc** (50.9F), **definefriends** (50.3M)

### Wrong

n = 19

- **Strongly predicted incorrect gender (80.0% of respondents and above)**
  - **johncenarises** (95.6M), **1-800-JustTheTipp** (93.7M), **-davros** (92.5M), **Captain-Mayhem** (88.3M)
- **Moderately predicted incorrect gender (60.0-79.9% of respondents)**
  - **Woodsbear** (79.3M), **TheHatredburritto** (78.8M), **find26** (78.5M), **Monkey1222** (73.3M), **freekythefruuk** (70.6M), **Cantthinkofonelol** (67.6M), **raeloxt** (67.1F), **Adorvex** (66.4M), **YaourtSurprise** (64.7M), **I\_eggo** (62.5M)
- **Slightly predicted incorrect gender (51.1-59.9% of respondents)**
  - **mentallydeteriorated** (59.8M), **Genshai** (58.6M), **-Fraddles-** (53.9M), **\_Onyx\_** (52.2M), **unbridledirony** (51.5F)

## ***Final thoughts***

### Trends

As you can see, respondents did much worse at predicting when a username was female. I think this has to do with the idea that Reddit is thought to be overwhelmingly male. This assumption appears to be so prevalent that it was visible even though the survey actually featured significantly more female usernames! I also found it interesting that all 4 usernames featuring “cat” in some way were guessed as belonging to a female and actually did belong to a female. What other patterns/trends stood out to you?

### Possible Improvements

- Some people, anticipating the bias discussed above, recommended I **include a “don’t know” option** instead of just male and female. I think this would have reduced the bias but at the expense of making too many people choose neither gender, which is something people fall back on when given the option even if they are actually leaning one way.
- Others suggested I include **options for how strongly the respondent felt about their answer**. 4 choices would probably be best here: male, leaning toward male, female, leaning toward female. That way, people would have to set their biases aside, at least more so than when being forced to choose between a binary. It would still give me the results I wanted in that people wouldn’t fall back on the “neither” option. Maybe someday I’ll conduct this type of survey with this 4-choice method. I’d use the same data set to sample from (all new usernames of course).
- One or two people suggested that I **let people know beforehand how many usernames of each gender were featured**. This an interesting suggestion, but I’m afraid it would make people super focused on making sure they got their numbers exactly right and make things more time-consuming.

I hope you found these results interesting! Let me know what you think.