PNAS RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES OPEN ACCESS

Does counting change what counts? Quantification fixation biases decision-making

Linda W. Chang[a,1], Erika L. Kirgios[b], Sendhil Mullainathan[c,d], and Katherine L. Milkman[a]

Affiliations are included on p. 11.

Edited by Timothy Wilson, University of Virginia, Charlottesville, VA; received January 22, 2024; accepted September 13, 2024

People often rely on numeric metrics to make decisions and form judgments. Numbers can be difficult to process, leading to their underutilization, but they are also uniquely suited to making comparisons. Do people decide differently when some dimensions of a choice are quantified and others are not? We explore this question across 21 preregistered experiments (8 in the main text, N = 9,303; 13 in supplement, N = 13,936) involving managerial, policy, and consumer decisions. Participants face choices that involve tradeoffs (e.g., choosing between employees, one of whom has a higher likelihood of advancement but lower likelihood of retention), and we randomize which dimension of each tradeoff is presented numerically and which is presented qualitatively (using verbal estimates, discrete visualizations, or continuous visualizations). We show that people systematically shift their preferences toward options that dominate on tradeoff dimensions conveyed numerically—a pattern we dub "quantification fixation." Further, we show that quantification fixation has financial consequences—it emerges in incentive-compatible hiring tasks and in charitable donation decisions. We identify one key mechanism that underlies quantification fixation and moderates its strength: When making comparative judgments, which are essential to tradeoff decisions, numeric information is more fluent than non-numeric information. Our findings suggest that when we count, we change what counts.

quantification | decision-making | numeracy | comparison fluency

Significance

Across 21 experiments with over 23,000 participants in managerial, policy, and consumer contexts, we identify a critical distortion that shapes how people make decisions involving tradeoffs across qualitative and quantitative attributes. When making hiring, donation, and policy decisions, people tend to privilege quantitative information, favoring options that dominate on the dimension described numerically. This "quantification fixation" is driven by the perception that numbers are easier to use for comparative decision-making; people who are more comfortable with numbers—those higher in subjective numeracy—are more likely to exhibit quantification fixation. As quantification becomes increasingly prevalent, the comparison fluency of numbers may systematically skew decisions. These findings suggest that quantifying certain choice features can have important repercussions for how decisions are made.

Quantification is spreading. It is reaching even the most intensely personal and ineffable parts of our lives. A proud new parent is not only given a newborn swaddled in cotton but also an Apgar score—a number from 0 to 10 that measures their baby's appearance, pulse, grimace response, activity, and respiration. Even hedonic experiences are quantified—the pleasure you can expect from drinking a beer is now distilled into a number that gives a sense of its bitterness (International Bitterness Units). Likewise, it is nearly impossible to discuss sports without statistics—batting averages are merely one of dozens of ways to value a baseball player; our personal favorite metric is NERD (a quantitative measure of a player's expected aesthetic value). Everywhere you turn there is a new rating or score turning an experience into a number.

The story of the 21st century may not just be about computers and the information age, but about the migration from the qualitative to the quantitative. This migration raises a fundamental question about the way we make judgments and decisions: Do we decide differently when some dimensions of a choice are quantified and others are not?

Psychology suggests a natural answer: Numbers are pallid and hard to process (1, 2) compared to the vividness of stories and richness of words (3, 4). The story of a single victim will draw more empathy than statistics describing the plight of multitudes. Moreover, some people—those low in numeracy—find it difficult to work with numbers, which hampers their judgments of and reliance on quantified information (5–8). In light of these findings, numeric information could, plausibly, be *disadvantaged*.

We argue for a second perspective on numbers when it comes to making choices—that quantification can create a psychological *advantage*. Choices require comparisons and tradeoffs. How do we decide between two job candidates when one has a higher grade point average but the other has more relevant work experience?

Quantification will powerfully sway us here because numbers are exactly suited for comparison: They allow us to judge what is larger and by how much. We can subtract, add, and divide—all operations used in making tradeoffs. In fact, when making choices, we tend to focus less on the absolute values of choice attributes, and more on their differences—which are easily and intuitively represented numerically (e.g., ref. 9). We may go so far as to mindlessly perform these mathematical operations when a problem does not require that we do so (10). As a result, people tend to value even superfluous metrics in decision-making,

Author contributions: L.W.C., E.L.K., S.M., and K.L.M. designed research; L.W.C. performed research; L.W.C. analyzed data; and L.W.C., E.L.K., S.M., and K.L.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2024 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

[1]To whom correspondence may be addressed. Email: changlw@sas.upenn.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2400215121/-/DCSupplemental.

Published October 28, 2024.

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

overlooking their irrelevance and failing to dwell on what a number does not represent (11, 12).

When making decisions that require tradeoffs, this suggests that people may experience greater *comparison fluency* when evaluating choice dimensions that are presented as numbers than when evaluating other choice attributes (13–16). We define comparison fluency as the felt and actual ease of using information to make judgments about differences. Moreover, we suggest that people are likely to be more sensitive to the magnitude of differences between numeric attribute values because numbers are more comparison fluent than non-numbers. As people become more sensitive to differences across attribute values, they also become more likely to rely on those attributes when choosing (17, 18). Thus, we propose that when deciding in the presence of tradeoffs, people place more weight on quantified choice attributes.

In this paper, we propose and show that decision-makers facing tradeoffs favor options that dominate on the numeric dimension—a phenomenon we refer to as *quantification fixation*, and we further show that the greater comparison fluency of numbers contributes to this pattern. Quantification fixation means that decision-makers, for example, may choose the job candidate with the higher grade point average but less relevant work experience even if, in principle, they care just as much about experience, simply because grade point average is quantified. Similarly, when deciding between restaurants, diners may choose the less expensive but noisier restaurant, simply because the average cost of a main course is often quantified via online review platforms like Google reviews while the ambiance is described verbally. And when selecting a health insurance plan, employees may opt for the one with the lower monthly premium even if it has a more limited network of providers, simply because the premium is presented numerically. We find evidence of quantification fixation in people's decisions in managerial contexts (i.e., a promotion decision, a conference location choice, a hiring decision), policy contexts (i.e., a decision about which public works project to pursue), consumer contexts (i.e., a hotel choice, a restaurant choice, a car choice, a property choice, a charity donation choice), and in incentive-compatible choices, including in the field (i.e., a personnel selection decision, a charity donation decision). Across all of the experiments we present, we ask participants to evaluate options involving tradeoffs across multiple attributes (e.g., price vs. quality) and select their preferred choice. We randomly vary which attribute is represented numerically (vs. qualitatively, using verbal estimates, discrete visualizations, or continuous visualizations). And we seek to hold the information conveyed about each choice constant irrespective of quantification. We consistently find that participants exhibit preference shifts in favor of whichever option is more attractive on the attribute described numerically.

Our findings are closely related to past research on choice evaluability (19, 20). The existing literature, however, has focused on evaluability based on objective criteria (e.g., how much information is provided), and here, we point out that there are subjective criteria (i.e., comparison fluency) that matter to evaluability as well. Notably, in our experiments, objective dimensions of evaluability are held constant. But numbers create a subjective ease of evaluability that is not captured by the previously identified objective features of evaluability. Formally, evaluability has previously been defined as the extent to which a decision-maker has the necessary reference information to judge attribute values (i.e., the extent to which something *can be* evaluated), and past theorizing has identified three factors that impact evaluability: mode (whether choice attributes are jointly or separately evaluated), knowledge (the amount of information available about the distribution of possible values each choice attribute could take on), and nature (whether an innate reference system exists for assessing the values of each choice attribute; for example, we know whether a particular temperature feels comfortable or not without being taught). Our experiments are designed to hold the objective evaluability of choice dimensions constant: Across experimental conditions, we consistently observe quantification fixation even when we (1) present options and attributes jointly, (2) share equivalent information about the same attributes of each option, and (3) provide attributes that do not have innate reference systems. And yet, even when choice attributes are equally evaluable on objective dimensions (based on mode, knowledge, and nature), we show that rendering them more subjectively evaluable by quantifying them and increasing their comparison fluency—i.e., making attributes more intuitively interpretable—skews decisions. Thus, quantification fixation is not driven by whether it is *possible* to evaluate attributes, but rather by whether it is *easy* to do so. Our findings suggest that comparison fluency should be recognized as another component of evaluability.

We show that quantification fixation emerges robustly across 21 preregistered experiments (20 of which demonstrate the effect, and 1 of which rules out a potential mechanism; see Table 1 for a summary of experiments in the main manuscript, total N = 9,303; see *SI Appendix*, Table S2 for a summary of supplemental experiments, total N = 13,936). Consistent with our theorizing, we also find evidence that quantification fixation is moderated by the comparison fluency of numeric information: Quantification fixation is attenuated when the numbers presented are harder to process (*Experiment 4* and *SI Appendix, Experiment S8*; see *SI Appendix*, Table S1 for an overview of tested mechanisms). Moreover, quantification fixation is mediated by people's perception that quantified information is more fluent and facilitates more fluent comparisons than qualitative information (*SI Appendix, Experiments S9a and S9b*). Decision-makers feel more confident and comfortable relying on numbers, and this feeling mediates the degree of quantification fixation we observe. Further, individuals with lower subjective numeracy are not as susceptible to quantification fixation (*Experiment 5*).

Our psychological perspective on numbers suggests an important force neglected in the spread of quantification. The act of quantification is not neutral. When we quantify a choice attribute (as is often done in settings ranging from Amazon to Glassdoor to Yelp), we systematically change decisions. The quantified dimension holds greater sway, while the qualitative is underweighted. In short, when we count, we change what counts.

Experiment 1a: Does Quantification Fixation Exist?

In *Experiment 1a* (N = 1,000, Amazon Mechanical Turk), we presented online participants with a hypothetical choice between two hotels that required them to make a tradeoff between price and rating. We randomized whether each hotel's price or rating was described numerically (vs. pictorially), emulating the type of information consumers might encounter on websites like Expedia. We predicted that participant choices would shift in favor of the hotel that dominated on whichever dimension was (randomly) presented numerically.

Results. We asked each participant to imagine that they were planning a vacation with their partner, who had been browsing a website with hotel listings and had identified two hotel options. Because of a discount expiring that night, participants were told they had to decide between the two hotels now. They were told that hotels on this website could vary in price from $100 to $500 per night and in rating from 1 to 5 stars, so knowledge about

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

Table 1. Overview of experiments

Exp	N	Choice context	Tradeoff	What does this study demonstrate?
1a	1k	Hotel choice	Price vs. ratings	Quantification fixation shifts decisions
1b	1k	Summer internship candidate choice	Calculus grade vs. management grade	Replication in a new context where there is similar familiarity with qualitative and quantitative information
1c	1k	Conference location choice	Connectedness vs. Sustainability	Replication in a new context; additionally, qualitative and quantitative descriptions are transparently linked
2	2k	Employee promotion choice	Likelihood of retention vs. Likelihood of advancement	Quantification fixation distorts preferences, shifting choice relative to baseline conditions where either (1) both dimensions are presented verbally or (2) both dimensions are presented numerically
3a	1k	Job candidate choice	Math Game performance vs. Angles Game performance	Replication in a new context with real financial incentives
3b	701	Charity donation choice	Accountability and Finance vs. Culture and Community	Replication in a new context with real donation decisions and in-person participants
4	2k	Public works project choice	Benefit vs. Efficiency	Quantification fixation is moderated by the fluency of quantified information
5	602	Charity donation choice	Accountability and Finance vs. Culture and Community	Replication in a nationally representative US sample making real donation decisions; additionally, the effect is moderated by subjective numeracy but not objective numeracy

the range of possible values for each attribute was held constant across conditions.

All participants evaluated the same two hotels and chose one to book. One hotel had a higher rating and higher price (i.e., "Hotel Luxe") and the other had a lower rating and lower price (i.e., "Hotel Milton").

Participants were randomly assigned to one of two conditions: *rating quantified* (n = 500) or *price quantified* (n = 500), which determined whether each hotel's rating or price was described numerically, with the other attribute described pictorially (as shown in Fig. 1*A*). Pictorial depictions were icon bars that included either five cash symbols or five stars. Given the price and rating ranges provided to participants, they could assume each cash symbol represented $100 and each star represented a 1.0 rating increment. After selecting a hotel, all participants answered follow-up questions about the experimental stimuli* and their demographics.

Consistent with our hypothesis, participants were significantly more likely to choose Hotel Luxe (the hotel with the higher rating and higher price), in the *ratings quantified* condition (51.6%) than the *price quantified* condition (33.0%), $\chi^2(1) = 34.68$, $P < 0.001$, 95% CI [0.124, 0.248], effect size $h = 0.379$ (see Fig. 2 for depiction of results).

To examine whether quantification fixation occurs in separate evaluation settings (vs. joint evaluation settings as in *Experiment 1a*), we ran another experiment (*SI Appendix, Experiment S1*) using this same hotel choice context. Here, however, we varied whether participants were presented with only (1) Hotel Luxe or (2) Hotel Milton and, separately, whether the hotel's (a) rating or (b) price was quantified, resulting in a total of four between-subjects conditions. Instead of deciding *which* hotel to book, participants were presented with a single hotel and decided *whether* they would book it, or not. As predicted and confirming that our findings replicate under separate evaluation, we again found evidence of quantification fixation: Participants were more likely to book the higher rated but more expensive hotel when rating (rather than price) was quantified, as evidenced by a significant interaction between assignment to the *rating quantified* conditions and assignment to the *higher rated more expensive hotel* (Hotel Luxe) conditions, $b_interaction = 0.281$, $SE = 0.062$, 95% CI [0.160, 0.401], $t(996) = 4.56$, $P < 0.001$.

To further establish the robustness of quantification fixation, we replicated this effect in a similar, preregistered joint evaluation experiment (*SI Appendix, Experiment S3b*) where participants chose between two restaurants and faced a tradeoff between cost versus commute time. Participants were more likely to choose the cheaper restaurant requiring a longer commute when price was quantified and distance was described with icon graphs than when distance was quantified and price was described pictorially, $\chi^2(1) = 83.28$, $P < 0.001$, 95% CI [−0.315, −0.205], effect size $h = 0.598$.†

While these initial studies were designed to simply demonstrate quantification fixation rather than to explore our proposed mechanism, we conducted a follow-up experiment to rule out the possibility that quantification provides a signal of importance (21–24). For example, people might infer that if someone took the time to generate numeric estimates to describe an attribute (rather than relying on verbal descriptions or pictorial visualizations), they did so because that (quantified) attribute is of particular significance. To address this, in another experiment (*SI Appendix, Experiment S3c*) we prompted participants to assess the importance of each attribute in the aforementioned restaurant choice scenario (*SI Appendix, Experiment S3b*). Participants were just as likely to indicate that cost was the more important attribute whether they were randomly assigned to see a restaurant's cost quantified (72.4%) or commute time quantified (76.4%), $\chi^2(1) = 0.851$, $P = 0.356$, 95% CI [−0.040, 0.120], effect size $h = 0.092$. An equivalence test using the two one-sided t test method shows that the outcomes of these two conditions are equivalent at 90% confidence, using a tolerance margin of ±0.105. This suggests it is unlikely that quantification fixation is

*See *SI Appendix* for more information about participants' perceived numeric values of qualitative stimuli. There, we report median perceived values of non-numeric stimuli across all experiments, as well as one-sample sign tests and dominance statistics testing for the equivalence of non-numeric and numeric stimuli. In some of our supplemental experiments, non-numeric stimuli were selected based on the median perceived numeric values of these stimuli in calibration pilots (that we describe where relevant, e.g., *SI Appendix*, Footnote 1). While participants' self-reported perceived values of non-numeric stimuli in our preregistered experiments and their corresponding numeric values in the quantified experimental condition do not match perfectly, the small inconsistencies are in different directions across experiments (i.e., sometimes the non-numeric information is perceived by the median participant to be slightly smaller and sometimes slightly larger than the numeric information). This variety makes it implausible that directional misperceptions of our non-numeric stimuli could explain our results.

†In another preregistered experiment (*SI Appendix, Experiment S3a*) where participants faced tradeoffs between annual property tax and distance to the city center when choosing a property, we again found evidence of quantification fixation, $\chi^2(1) = 31.4$, $P < 0.001$, 95% CI [−0.208, −0.100], effect size $h = 0.363$.

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

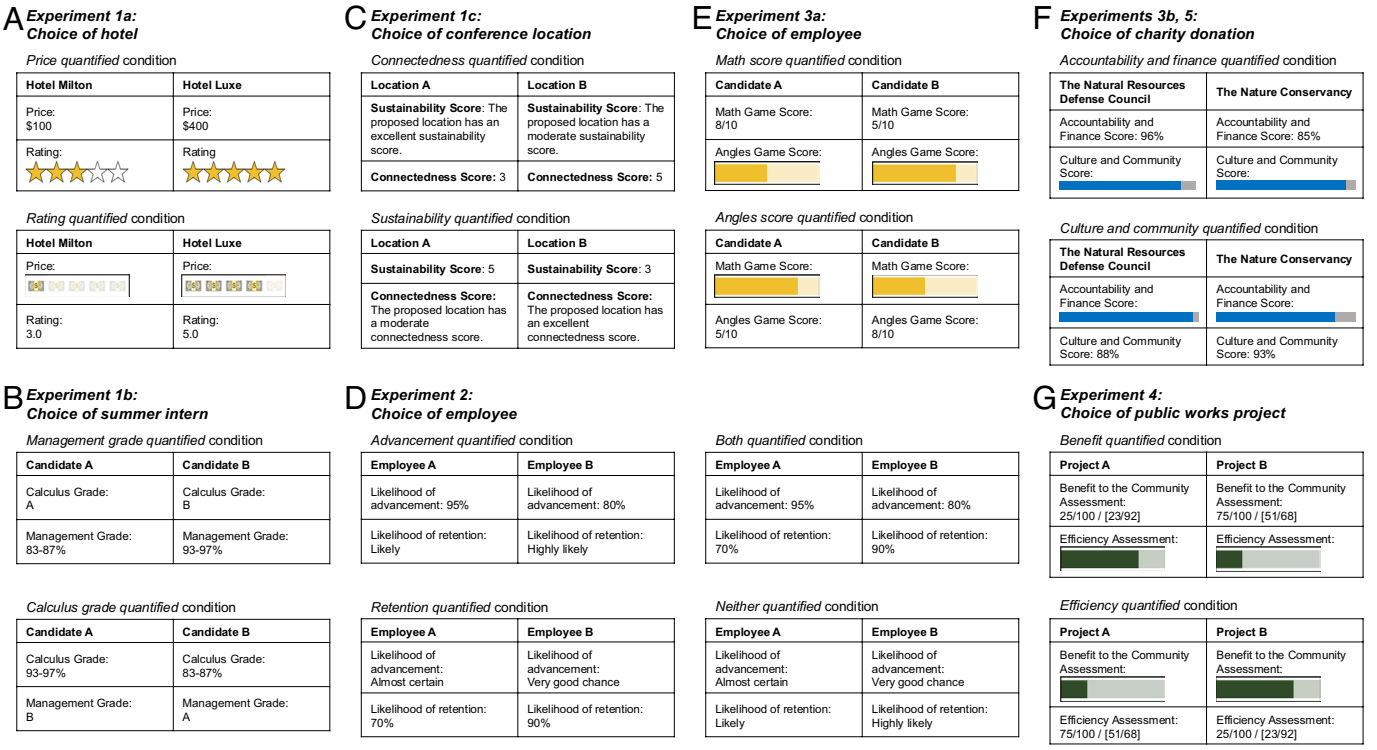

Fig. 1. Overview of experimental stimuli and experimental conditions. Key experimental stimuli showing choice tradeoffs from (*A*) *Experiment 1a* where participants saw either hotels' *price quantified* or hotels' *rating quantified*. (*B*) *Experiment 1b* where participants saw either candidates' *calculus grade quantified* or candidates' *management grade quantified*. (*C*) *Experiment 1c* where participants saw either conference locations' *connectedness quantified* or locations' *sustainability quantified*. (*D*) *Experiment 2* where participants saw employees' *likelihood of advancement quantified, likelihood of retention quantified, both quantified,* or *neither quantified*. (*E*) *Experiment 3a* where participants saw either employees' *math score quantified* or their *angles score quantified*. (*F*) *Experiments 3b and 5* where participants saw either charities' *accountability and finance quantified* or charities' *culture and community quantified.* (*G*) *Experiment 4* where participants saw either public works projects' *benefit quantified* or projects' *efficiency quantified*, and numeric scores that were either relatively *disfluent* (in brackets) or *fluent*.

driven by shifts in attributes' perceived importance when they are quantified, as we detect no such shift.[‡]

All experiments described up to this point have relied on icons to visualize non-numeric information, which may confound quantification and scale expansion given that the difference between, say, $100 and $400 may feel bigger than the difference between $ and $$$$ (25). However, in the experiments that follow we rely on different presentations of non-numeric information to establish the robustness of our effect, and in our next experiment, we examine whether quantification fixation persists when attributes are similarly familiar when described numerically or non-numerically.

Experiment 1b: Does Quantification Fixation Persist When Numeric and Non-Numeric Descriptors Are Both Familiar?

In *Experiment 1b* (N = 1,000, Prolific), we examined whether quantification fixation persists when participants are expected to be very familiar with an attribute's meaning whether it is described numerically or non-numerically (in this case, grades presented as numbers or letters).

Results. We asked participants to imagine that they were deciding between two candidates to hire for a competitive summer internship. The application asked candidates to report two relevant pieces of coursework: their most recent grade in a math class and their most recent grade in a business class. Participants were told that these two candidates had performed similarly in their interviews and had

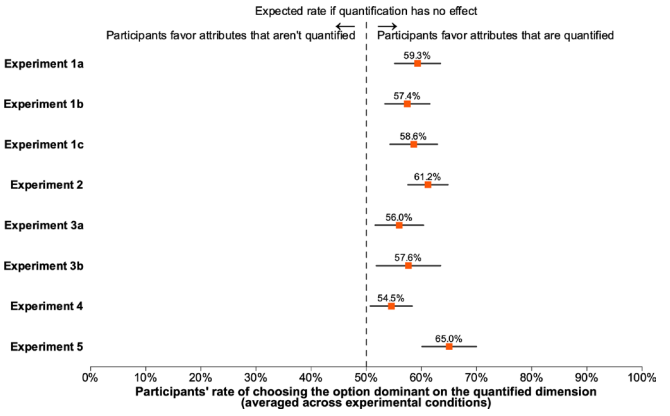

Fig. 2. Across eight preregistered experiments, people privilege numeric information when making tradeoffs, favoring options that dominate on the quantified dimension. Across all experiments, participants evaluate two options that involve a tradeoff across two attributes and select their preferred choice. We randomly vary which attribute is quantified across experimental conditions. Without quantification fixation, the probability of choosing the option that dominates on the quantified dimension should be 50%, averaged across conditions (dashed line). The orange points depict the likelihood of choosing the option dominant on the quantified dimension (i.e., the effect of quantification fixation on choice). In each of eight preregistered experiments, the likelihood of choosing the option that dominates on the quantified dimension significantly exceeds 50%. All error bars correspond to 95% CI.

[‡]In our *SI Appendix*, we also report the results of an experiment using icon stimuli that sought to probe another potential mechanism for quantification fixation. Specifically, in *SI Appendix, Experiment S2*, we sought to examine whether people weight quantified attributes more heavily because numeric information is encoded and recalled better than non-numeric information. Participants faced the same hotel choice as in *Experiment 1a* and then were asked to recall the price and rating of each hotel they evaluated. Participants were marginally more likely to accurately recall an attribute value when it was quantified (93.3%) as compared to when it was not (91.7%), *b_Quantified* = 0.016, SE = 0.009, 95% CI [−0.0003, 0.032], t(996) = 1.78, *P* = 0.075. However, differences in recall rates across dimensions did not mediate quantification fixation (*b_recallDiff* = 0.005, *SE* = 0.059, 95% CI [−0.104, 0.114], t(998) = 0.087, *P* = 0.931), providing only weak evidence for this mechanism.

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

excellent application materials. All participants decided between the same two candidates, one with a higher calculus grade but lower management grade and the other with a lower calculus grade but higher management grade.

Participants were randomly assigned to one of two conditions: *calculus grade quantified* (n = 499) or *management grade quantified* (n = 501), which determined whether candidates' calculus grade or management grade was described as a numeric range (e.g., 93 to 97%), with their other grade described by a letter (e.g., "A," as shown in Fig. 1*B*).

Replicating the results of our prior experiments, we find that participants were significantly more likely to choose the candidate with the higher management grade in the *management grade quantified* condition (83.8%) than the *calculus grade quantified* condition (68.9%), $\chi^2(1) = 29.94$, $P < 0.001$, 95% CI [0.095, 0.203], effect size $h = 0.355$ (see Fig. 2 for depiction of results). These results suggest that even in a setting where the numeric and non-numeric information provided about a choice are similarly familiar, quantification fixation persists. They also demonstrate that our results hold even when information about choices is provided in the form of an imprecise numeric range (e.g., a grade of 93 to 97%), suggesting our findings are not driven by how precise numeric information feels.

In *SI Appendix, Experiment S4*, we directly tested the impact of making the numbers used to describe choice attributes less precise. We did this by varying whether we presented participants with a range of values (i.e., an estimated price of $25 to $45) or a single point estimate (i.e., an estimated price of $35) to describe one dimension of a tradeoff between restaurants. Quantification fixation persisted regardless of whether a range or point estimate was provided. Taken together, these results suggest that the perceived precision of numeric (vs. non-numeric) information is not a key driver of quantification fixation.

Experiment 1c: Does Quantification Fixation Persist When Mappings from Qualitative to Quantitative Scores Are Transparent?

In *Experiment 1c* (N = 1,000, Prolific), we assessed whether quantification fixation is robust to presenting qualitative attributes verbally and with direct translations to their quantitative equivalents.

Results. We asked participants to imagine that they were a manager at a financial firm that was planning to host a conference, and their job was to choose a conference location from two potential sites that were similar in cost. Participants learned that a committee had assessed and scored these two proposed locations on two dimensions: their connectedness and their sustainability (see *SI Appendix* for details). Scores from 1 to 5 were possible for both locations' connectedness and sustainability, and each numeric score was given a corresponding verbal description that was presented to participants and available at the time they made their decision. This presentation eliminated any potential ambiguity about how to convert qualitative information into quantified information (and vice versa). For example, the sustainability score was described as follows:

The Sustainability Score is based on the environmental consequences of holding a large meeting in the proposed location.

- 5: The proposed location has an excellent sustainability score.
- 4: The proposed location has a good sustainability score.
- 3: The proposed location has a moderate sustainability score.
- 2: The proposed location has a fair sustainability score.
- 1: The proposed location has a poor sustainability score.

Participants were randomly assigned to one of two conditions: the *connectedness quantified* condition (n = 500) or the *sustainability quantified* condition (n = 500). In both conditions, participants evaluated the same two proposed conference locations and chose one; all that varied across conditions was which dimension—connectedness versus sustainability—was described numerically and which was described verbally (see Fig. 1*C*, for more details). One conference location had a higher connectedness score and a lower sustainability score while the other conference location had a lower connectedness score and a higher sustainability score.

Again, we found evidence of quantification fixation: Participants were significantly more likely to choose the conference location with the higher connectedness score but lower sustainability score in the *connectedness quantified* condition (78.0%) than in the *sustainability quantified* condition (60.8%), $\chi^2(1) = 34.02$, $P < 0.001$, 95% CI [0.114, 0.230], effect size $h = 0.377$ (see Fig. 2 for depiction of results). Here, we see that even when verbal descriptions are transparently mapped to numeric ratings, people continue to exhibit quantification fixation.

Experiment 2: Does Quantification Fixation Distort Preferences?

In *Experiment 2* (N = 2,000, Prolific), we sought to explore whether quantification fixation distorts preferences compared to contexts where all attributes are presented in the same format. As in our prior experiments, all participants made a decision involving a tradeoff between two choice attributes (here: an employee's potential vs. their loyalty). But we added new experimental conditions to assess people's preferences over the (employee) choice under consideration when all attributes were presented using identical formats. Specifically, in addition to our standard stimuli presenting just one choice attribute numerically, we included (1) a new condition in which *both* choice attributes were described numerically and (2) a new condition in which both choice attributes were described verbally. This design allowed us to assess the distortion of preferences produced by quantifying just one choice attribute.

Results. Participants imagined that they were a manager deciding which of two software engineers to promote. We told them that these engineers had been assessed on two promotion criteria: likelihood of advancement and likelihood of retention. We then explained these criteria in greater detail. In all conditions, participants decided between the same two employees: one with a higher likelihood of advancement but a lower likelihood of retention and the other with a lower likelihood of advancement but higher likelihood of retention.

Participants were randomly assigned to one of four conditions: *advancement quantified* (n = 500), *retention quantified* (n = 500), *both quantified* (n = 501), or *neither quantified* (n = 499). Condition assignment determined which attribute (retention, advancement, both, or neither) was presented as a number; nonquantified dimensions were presented verbally (see Fig. 1*D* for more details). Verbal stimuli were precisely matched to numeric stimuli based on results from a prior study on perceptions of the probabilities conveyed by different words (26). Moreover, as in our prior experiment (*Experiment 1c*), we transparently mapped verbal descriptors to numeric values (i.e., we explicitly told participants "Almost certain = 95%"; "Likely = 70%") to eliminate ambiguity.

To understand how quantification fixation skews preferences, we first examined participants' likelihood of choosing the employee with the higher likelihood of advancement but lower likelihood

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

of retention in our two benchmark conditions: the *both quantified* (27.9%) and *neither quantified* (32.7%) conditions. We found no significant difference in selection decisions between these two conditions (P = 0.104; see *SI Appendix*, Table S14 for regression results). That is, when both choice attributes were presented the same way, choices remained consistent. Quantifying just one attribute, however, shifted choices away from this preference. Specifically, compared to these benchmark conditions, participants were more likely to select the employee rated as more likely to advance through the ranks when advancement (but not retention) was quantified (44.2%, Ps < 0.001), and less likely to select this employee when retention (but not advancement) was quantified (21.8%, Ps < 0.05). Finally, we replicate our standard quantification fixation effect: Participants were more likely to select the employee with the higher likelihood of advancement in the *advancement quantified* condition than in the *retention quantified* condition, P < 0.001 (see Fig. 2 for depiction of these results).

Experiment 3a: Does Quantification Fixation Persist When Incentives Are on the Line?

In *Experiment 3a* (N = 1,000, Prolific), we sought to test whether our results extend to an incentive-compatible decision environment. In addition, we varied the way stimuli were presented. While choice features are frequently described verbally or with icons in real decision environments (e.g., on platforms like Yelp and Airbnb), there are also contexts where continuous ratings like bar graphs are used to convey information (e.g., on platforms like Charity Navigator, Glassdoor, and Consumer Reports). Here, we communicated about 0 to 10 "scores" describing different choice options and we conveyed those scores either numerically or with a bar graph filled in proportionally to the score (e.g., a score of 5 out of 10 would be represented by a bar graph that is 50% full). This ensured that the quantitative and qualitative information provided about choice attributes was equally granular.

Results. Participants were truthfully told that they would hire another Prolific worker to serve as their "employee" and that they would be paid based on their chosen employee's performance in a game. The candidates available for hire were all real Prolific workers who had previously completed three different games (the Math Game, Angles Game, and Trivia Game; see *Methods*), and in those games, they were paid for each correct response.

Participants read about each of the three games that candidates played and learned that they would be provided with information about candidates' performance on two of these games—the Math Game and the Angles Game. They also learned that they would earn a bonus based on their chosen employee's performance on a third game—the Trivia Game. Specifically, they would earn $0.05 for each trivia question their selected employee answered correctly. Participants viewed a pair of candidates' profiles and selected one of the two candidates to hire. In each candidate pairing, one candidate had a higher Math score and a lower Angles score while the other candidate had a lower Math score and a higher Angles score.

Participants were randomly assigned to one of two conditions: the *math score quantified* condition (n = 501) or the *angles score quantified* condition (n = 499), and this determined which score (the Math Game score or Angles Game score) was conveyed numerically and which was conveyed pictorially via a yellow bar graph that was either 40%, 50%, or 80% filled in to represent scores of 4/10, 5/10, and 8/10, respectively. Example stimuli are shown in Fig. 1*E*. In order to increase generalizability, we stimulus-sampled and presented participants with one of two different Prolific worker profile pairings.[§]

Replicating prior results, participants were significantly more likely to choose the candidate with the higher Math score but lower Angles score in the *math score quantified* condition (66.5%) than in the *angles score quantified* condition (54.5%), $\chi^2(1) = 14.46$, P < 0.001, 95% CI [−0.182, −0.057], effect size h = 0.245 (see Fig. 2 for depiction of results).

Math scores were actually more predictive of Trivia performance than Angles scores, and as a result, the "right" choice would have been for participants to weight math scores more heavily in their candidate selection decisions. However, due to quantification fixation, participants left more bonus payments on the table in the *angles score quantified* condition (*Mean bonus payments* = $0.26, SD = 0.10) than in the *math score quantified* condition (*Mean bonus payments* = $0.28, *SD* = 0.10), $t(996.41) = 3.49$, P < 0.001, earning 6% less in the *angles score quantified* condition.[¶]

This experiment confirms that quantification fixation arises even when cash rewards are on the line.[#]

Experiment 3b: Does Quantification Fixation Persist in Real, In-Person Decisions?

In *Experiment 3b*, we tested for quantification fixation in an experiment involving real, in-person donation decisions.

Results. Participants in *Experiment 3b* (N = 701, in-person data collection from three sites: two university campus labs and one local Chicago pop-up lab/storefront, Mindworks) learned that they would choose a charity to receive a $1 donation. They were (truthfully) informed that the two candidate charities presented were assessed by an independent auditor, Charity Navigator, on multiple dimensions. Those dimensions were the charity's Accountability and Finance score and the charity's Culture and Community score. Participants were given additional information about the criterion for these scores and about each charity's mission (see *SI Appendix* for more information). One charity, The Natural Resources Defense Fund, had a higher Accountability and Finance score and a lower Culture and Community score. The other charity, The Nature Conservancy, had a lower Accountability and Finance score and a higher Culture and Community score.

Participants were randomly assigned to one of two conditions: the *accountability and finance quantified* condition (n = 349) or the *culture and community quantified* condition (n = 352). As in prior experiments, all that varied across conditions was which dimension was described numerically and which was described with a filled bar graph (as shown in Fig. 1*F*).

[§]The game score tradeoff for the first pairing was three points: The higher Math Score but lower Angles Score candidate had a Math Game score of 8/10 and an Angles Game score of 5/10, while these values were inverted for the lower Math Score but higher Angles Score candidate. The game score tradeoff for the second pairing was four points: the higher Math Score but lower Angles Score candidate had a Math Game score of 8/10 and an Angles Game score of 4/10 while these scores were swapped for the lower Math Score but higher Angles Score candidate.

[¶]To establish the robustness of quantification fixation with continuous visualizations of the qualitative attribute, we replicated this effect in two supplemental experiments where participants weighed tradeoffs: (1) a preregistered car choice scenario experiment (*SI Appendix, Experiment S5a*) and (2) a preregistered public works project choice scenario experiment with bar graphs that included tick marks to facilitate the transparent mapping of bar graph values to numeric scores and a 10-second delay during scenario evaluation to encourage participants to carefully evaluate options (*SI Appendix, Experiment S5b*). In both experiments, we varied across conditions which dimension of choice was represented numerically and which was represented via a continuous visualization. We found additional, robust evidence of quantification fixation in these two supplemental experiments (P's<0.001).

[#]The rewards on offer here were large enough to double participants' base pay.

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

Participants exhibited quantification fixation, donating significantly more often to the charity with the higher Accountability and Finance score (The Natural Resources Defense Fund) when the Accountability and Finance score was quantified (56.7%) than when the Culture and Community score was quantified (41.4%), *b _ AccountabilityFinanceQuantified* = 0.153, *SE* = 0.037, 95% CI [0.080, 0.226], $t(697) = 4.09$, $P < 0.001$ (see Fig. 2 for depiction of results; see *SI Appendix*, Table S20 for full regression results).

We also ran a field replication of this experiment (*SI Appendix, Experiment S7*) by posting an ad across Facebook and Instagram using Meta's advertising software. The ad urged users to click on a link to vote for one of two environmental charities to receive a $1,000 donation. Upon clicking the link, social media users faced the same choice between two charities as participants in the aforementioned study. We struggled to recruit our targeted sample of 1,000 users, only attracting 236 participants in the experiment's preregistered maximum duration to reach its target sample size. Despite this challenge, we had sufficient power to detect that participants were marginally more likely to vote for the charity with the higher Accountability and Finance score but lower Culture and Community score in the *accountability and finance quantified* condition (48.3%) than in the *culture and community quantified* condition (35.6%), $\chi^2(1) = 3.41$, $P = 0.065$, 95% CI [−0.006, 0.260], effect size $h = 0.258$, an effect size that was comparable to those observed in prior studies. Taken together, *Experiments 3a and 3b* and *SI Appendix, Experiment S7* suggest that quantification fixation holds when incentives are on the line, even when people are making decisions in the wild.

Experiment 4: Does the Comparison Fluency of Numeric Information Moderate Quantification Fixation?

In *Experiment 4* (N = 2,000, Amazon Mechanical Turk), we explored a key mechanism that we hypothesize contributes to quantification fixation: the comparison fluency of numeric information (i.e., the ease with which numeric information can be compared). We predicted that when numeric information was presented in a way that reduced the ease of comparison, quantification fixation would be attenuated.

Results. We asked each participant to imagine they lived in a city where community members vote on how to spend part of a public budget. Their job was to help select which of two potential projects the city's budget should fund. Both project proposals involved building a park in a neighborhood that did not have one, and a team of volunteer budget delegates had purportedly assessed both proposals on two dimensions: the project's expected benefit to the community and the project's expected efficiency. Participants received detailed information about how each score was determined (see *SI Appendix* for details). One proposal had a higher assessed benefit to community score and a lower efficiency score and the other had a lower assessed benefit to community score and a higher efficiency score.

We randomly assigned participants to either the *benefit quantified* or *efficiency quantified* conditions, which determined which attribute was presented as a number (vs. a bar graph). We also introduced a new manipulation to our experimental design (see Fig. 1*G* for more details): Some participants were assigned to a *fluent number* condition and were shown numeric scores that were easy to process (e.g., 75/100 or 25/100) while others were assigned to a *disfluent number* condition and were shown numeric scores that were more difficult to process (e.g., 51/68 and 23/92).[‖] This yielded four experimental conditions in a 2 × 2 between-subjects design (*benefit quantified-fluent number:* n = 498, *benefit quantified-disfluent number:* n = 501, *efficiency quantified-fluent number*: n = 500, *efficiency quantified-disfluent number:* n = 500).

Consistent with our theory that comparison fluency drives quantification fixation, we found a significant interaction between assignment to the *benefit quantified* conditions and assignment to the *disfluent number* conditions, *b_BenefitQuantified-x-DisfluentNumber* = −0.148, *SE* = 0.037, 95% CI [−0.221, −0.075], $t(1996) = -3.99$, $P < 0.001$. Quantification fixation was attenuated when the numeric information was presented in a way that made comparisons less fluent (see Fig. 3*A* for depiction of results).

As a result of this significant interaction, we did not find a main effect of assignment to the *benefit quantified* conditions, *b_BenefitQuantified* = 0.017, *SE* = 0.019, 95% CI [−0.020, 0.053], t(1996) = 0.899, $P = 0.369$. We also did not find (nor did we predict) a main effect of assignment to the *disfluent number* conditions on participants' likelihood of choosing the higher benefit but less efficient project proposal, *b_DisfluentNumber* = 0.034, *SE* = 0.027, 95% CI [−0.018, 0.085], $t(1996) = 1.25$, $P = 0.210$. This experiment provides suggestive evidence that manipulating the comparison fluency of numeric information can attenuate quantification fixation such that people are less likely to overweight numeric information that is more difficult to use when making comparative judgments.

To establish the robustness of this finding, we replicated this effect ($P = 0.025$) in a similar, preregistered experiment using the same scenario with different numeric values assigned to choice attributes, (e.g., 90/100 instead of 75/100; see *SI Appendix, Experiment S8* for complete study details). In two further experiments (*SI Appendix, Experiments S9a and S9b*), we provide convergent evidence for comparison fluency as a mechanism: We show that quantification fixation is partially mediated by a three-item measure of fluency (measuring comfort, confidence, and ease with using numeric vs. non-numeric information).

Experiment 5: Do Objective and Subjective Numeracy Impact Susceptibility to Quantification Fixation?

In *Experiment 5*, we sought to test whether quantification fixation would replicate in a nationally representative sample of adults from the United States making incentive-compatible decisions about where to donate money. Moreover, we sought more evidence that comparison fluency drives quantification fixation by measuring participants' subjective and objective numeracy as potential moderators of quantification fixation.

Why might subjective or objective numeracy moderate quantification fixation, and what would this imply about the mechanism underlying the effect? To the extent that decision-makers feel more comfortable and confident using numbers to make comparisons, they may overrely on numeric information. Meanwhile, people who find numeric information more difficult to process may avoid using numeric information and instead rely on visual and verbal estimates. While objective numeracy measures people's ability to accurately work with mathematical concepts and numbers (27), subjective numeracy measures people's comfort using numbers (28). Indeed, recent work

[‖]To ensure the validity of this manipulation, we pretested 22 numeric stimuli with 220 participants who each rated five numbers (resulting in ~50 ratings per number) along 3 dimensions. Specifically they rated (on a 1 to 7 scale; Strongly disagree to Strongly agree) each number on whether it was (1) difficult to interpret, (2) weird, and (3) easy to use in calculations (reverse-coded). We averaged these ratings to create a composite where smaller numbers correspond to higher perceived comparison fluency and larger numbers correspond to higher perceived comparison disfluency, confirming that 51/68 and 23/92 are similarly disfluent ($M_{51/68} = 5.11$, $M_{23/92} = 4.93$), while their corresponding round numbers—75/100 and 25/100—are similarly fluent ($M_{75/100} = 2.10$, $M_{25/100} = 1.90$).

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

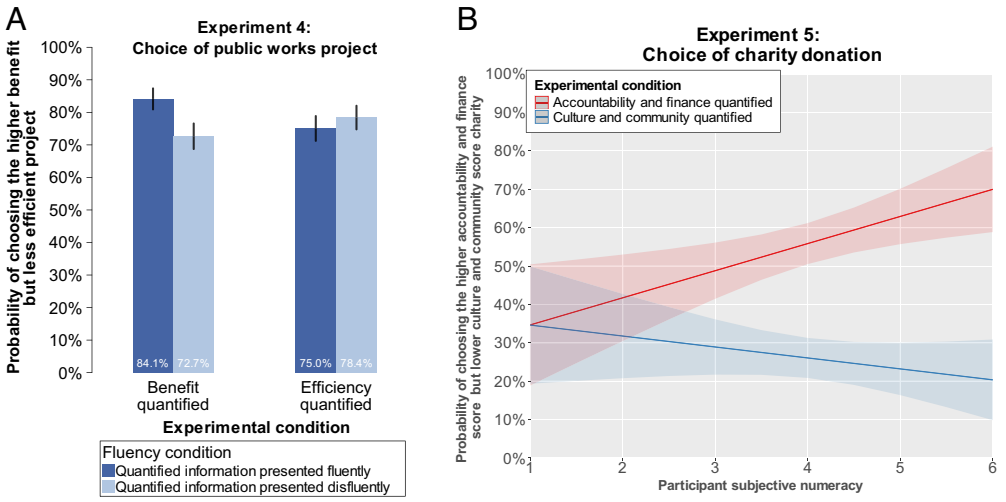

Fig. 3. Across two preregistered experiments, we identify comparison fluency of quantified information as a mechanism that contributes to quantification fixation. (*A*) Experiment 4: When quantified information is relatively disfluent, people are less susceptible to quantification fixation. (*B*) Experiment 5: People who report higher subjective numeracy (those who feel more comfortable or at ease with numbers) are more likely to display quantification fixation. All error bars correspond to 95% CI.

proposes that while objective numeracy measures a form of cognition (i.e., ability to accurately work with numbers), subjective numeracy measures a type of metacognition [i.e., ease with numbers; (29)]. If comparison fluency drives quantification fixation, then we would expect less subjectively numerate individuals to show this effect to a lesser extent. If objective ability to accurately work with numbers drives quantification fixation, then we would expect less objectively numerate individuals to show this distortion to a lesser extent as well. Therefore, in this experiment, we measured participants' subjective and objective numeracy and assessed whether and how these individual differences affected participants' susceptibility to quantification fixation.

Results. Participants (a nationally representative sample of 602 adults from the United States recruited via Qualtrics Panels) were first asked to provide information about their age, gender, race, geographic region, and education level (following the Qualtrics Panels team's recommended protocols). We then presented them with the charity choice scenario described in *Experiment 3b* in which participants faced a tradeoff between supporting a charity with a higher Accountability and Finance score but lower Culture and Community score (the Natural Resources Defense Fund) and a second charity with a higher Culture and Community score but lower Accountability and Finance score (The Nature Conservancy). As in *Experiment 3b*, participants were told that they would make a choice between these two charities and that we would donate $1.00 to the charity they selected.

Participants were randomly assigned to one of two conditions: the *accountability and finance quantified* condition (n = 293) or the *culture and community quantified* condition (n = 309), and as usual, condition assignment determined which score for each charity was described numerically (versus graphically, as shown in Fig. 1*F*). Participants chose one of the two charities as the recipient of their $1.00 donation and then answered follow-up questions about the experimental stimuli. Next, we measured participants' objective numeracy and subjective numeracy.

Replicating prior studies, we found evidence of quantification fixation in this nationally representative participant sample: Participants were significantly more likely to donate to The Natural Resources Defense Fund in the *accountability and finance quantified* condition (56.0%) than the *culture and community quantified* condition (25.9%), $\chi^2(1) = 55.22$, $P < 0.001$, 95% CI [0.223, 0.379], effect size $h = 0.623$ (see Fig. 3*B* for depiction of results).[**]

We next explored potential moderation by objective and subjective numeracy with three models (see Table 2, Models 1 to 3). In all models, we include an indicator for assignment to the *accountability and finance quantified* condition. In Model 1, we also include participants' mean-centered continuous score on the objective numeracy scale and an interaction between this variable and the indicator for assignment to the *accountability and finance quantified* condition. In Model 2, we instead include participants' mean-centered continuous score on the subjective numeracy scale and an interaction between this variable and the indicator for condition assignment. In Model 3, we include both participants' mean-centered score on the objective numeracy scale and the subjective numeracy scale, as well as both two-way interactions between the numeracy variables and the indicator for condition assignment.

Across all three models, we consistently see a strong main effect of assignment to the *accountability and finance quantified* condition ($P < 0.001$). In Model 1, we do not see any significant main effect of participants' objective numeracy or any interaction between participants' objective numeracy and the indicator for condition assignment. In contrast, in Model 2 we find a significant interaction between participants' subjective numeracy and the indicator for condition assignment ($P = 0.005$). Specifically, our results suggest that quantification fixation is greater for individuals who report higher levels of comfort working with numeric information; conversely, those who are less comfortable working with numbers are less likely to exhibit quantification fixation (see Fig. 3*B* for a depiction of these results).[††] Finally, in Model 3 we continue to find a significant interaction between participants' subjective numeracy and the indicator for condition assignment ($P = 0.010$) but no significant interaction between participants' objective numeracy and the indicator for condition assignment ($P = 0.842$).

In sum, while objective numeracy does not moderate quantification fixation, subjective numeracy moderates the effect of quantification fixation on choice. Those with higher subjective numeracy, who find numbers more fluent, are more susceptible to

[**]This result is robust to controlling for participant demographics (age, gender, race, geographic region, and education level), $P < 0.001$. See *SI Appendix* for more details.

[††]Consistent with our preregistration, we ran two more models examining the interaction of the indicator for assignment to the *accountability and finance quantified* condition and the ability subscale and preference subscale of the subjective numeracy scale separately. We found similar results for both (Ps < 0.05); self-reported ability and ease in working with numbers both independently predicted susceptibility to quantification fixation. These results provide additional support for our proposed mechanism of comparison fluency. See *SI Appendix* for further details.

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

Table 2. *Experiment 5* regression results

	Model 1		Model 2		Model 3	
	Estimate	*P*	Estimate	*P*	Estimate	*P*
Accountability and finance quantified condition	0.300*** (0.039)	<0.001	0.302*** (0.038)	<0.001	0.301*** (0.038)	<0.001
Participant's objective numeracy (mean-centered)	−0.032 (0.026)	0.222			−0.024 (0.028)	0.397
Accountability and finance quantified condition*Participant's objective numeracy	0.054 (0.041)	0.196			0.009 (0.045)	0.842
Participant's subjective numeracy (mean-centered)			−0.029 (0.023)	0.210	−0.021 (0.024)	0.378
Accountability and finance quantified condition*Participant's subjective numeracy			0.099** (0.035)	0.005	0.097* (0.038)	0.010
Intercept	0.260*** (0.025)	<0.001	0.259*** (0.025)	<0.001	0.260*** (0.025)	<0.001
Observations	602		602		602	
Adjusted R^2	0.092		0.103		0.102	

Note: This table reports the results of three ordinary least squares (OLS) regressions predicting whether the Natural Resources Defense Fund was chosen. The primary predictor in these regressions is an indicator for assignment to the *accountability and finance quantified* condition (the *culture and community quantified* condition is the omitted comparison group). Model 1 includes each participant's mean-centered objective numeracy and its interaction with the indicator for assignment to the *accountability and finance quantified* condition. Model 2 includes each participant's mean-centered subjective numeracy and its interaction with the indicator for assignment to the *accountability and finance quantified* condition. Model 3 includes each participant's mean-centered objective numeracy, each participant's mean-centered subjective numeracy, as well as both two-way interactions between the numeracy variables and the indicator for assignment to the *accountability and finance quantified* condition. SE reported in parentheses are estimated robustly using HC3. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

quantification fixation. In other words, a discomfort with numbers seems to shield people from exhibiting quantification fixation.

General Discussion

Daily life includes hundreds of decisions, big and small, that require people to make tradeoffs—for example, whether to drive to work or take public transportation, whether to take a higher-paying job or a more fulfilling one, and whether to allocate resources to a lower risk project or one that has more potential. Our research explores the impact of quantification on decisions that involve weighing competing attributes. These kinds of decisions are made billions of times each day and are facilitated by information aggregators, governments, and retailers ranging from Glassdoor to Healthcare.gov to Yelp to Airbnb. We identify a critical distortion that shapes the way we make tradeoffs across choice attributes (e.g., a health plan's quality vs. its copay).

Across eight preregistered experiments with over 9,000 participants (and 13 preregistered supplemental experiments with an additional ~13,900 participants), we find robust evidence of quantification fixation: When faced with decisions that involve tradeoffs across qualitative and quantitative attributes, people privilege the quantitative, favoring different job candidates, projects, and charities based on which option is more attractive on dimensions described numerically (see Fig. 2 for overview of quantification fixation across experiments). This effect persists across managerial, policy, and consumer choices in both joint and separate evaluation contexts, and it affects decisions that involve both objective and subjective tradeoffs. Importantly, we show that quantification fixation has financial consequences, emerging even when cash rewards are on the line (e.g., in an incentive-compatible hiring task) and when a real donation to one of two charities is at stake.

We find that the higher comparison fluency of numeric information contributes to quantification fixation. When numeric information is relatively disfluent, people are less susceptible to quantification fixation. Fluency also mediates the effect, and people who report higher subjective numeracy (indicating that they are more comfortable or fluent with numbers) are more likely to display quantification fixation. Together, these findings suggest that quantification fixation is driven by people's comparison fluency with numeric tradeoffs. It would be worthwhile for future research to confirm an implication of our theorizing: That by increasing the comparison fluency of non-numeric information, quantification fixation can be reduced. Relatedly, future work might explore whether people underweight information that does not feel comparison fluent when making tradeoffs; in other words, whether quantification fixation is also a form of nonquantification neglect.

Our findings suggest that evaluability theory (19, 20), which identifies three objective contributors to evaluability, may be incomplete. We propose that comparison fluency—the subjective ease with which a decision-maker can interpret and compare attribute values—is a critical fourth factor that affects evaluability and, consequently, decision-making outcomes, even when the feasibility of objectively evaluating attributes is held constant. Quantification fixation may be just one important distortion arising from variations in the fluency with which choice attributes can be evaluated.

Although we provide evidence that people's fluency with quantified information contributes to quantification fixation, the phenomenon is likely multiply determined. In fact, we find that numbers are encoded and recalled marginally more successfully than non-numbers, which may contribute to quantification fixation (though this effect is quite small and we find no evidence of mediation; see *SI Appendix, Experiment S2*). In *SI Appendix*, Table S1, we present an overview of alternative mechanisms for quantification fixation that we examined and ruled out across our experiments. For example, while an aversion to the perceived ambiguity of non-numeric information may contribute to quantification fixation (30, 31), we show that the effect arises even when we transparently map non-numeric information onto numbers (*Experiments 1c and 2*). We also find that attributes were perceived as similarly important whether they were quantified or not (*SI Appendix, Experiment S3c*), but additional research examining whether people perceive numeric information as more trustworthy (or more trusted by experts), driving overreliance on that information, would be informative (32). More broadly, it would be valuable for future work to probe the role of additional basic psychological mechanisms, such as an attribute's memorability and tendency to draw attention, that may contribute to quantification fixation.

Moreover, our experiments do not allow us to distinguish between the role of actual and felt ease in using numbers for comparative judgments. While we have taken care to ensure that people are equally able to assess differences across numeric and non-numeric attributes

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

in our experiments, it is more straightforward and often requires fewer steps to calculate differences between numbers than non-numbers (e.g., in *Experiment 1c*, participants might map verbal estimates back to numbers and then subtract the two numbers from each other—the first mapping step is not required when the information is immediately presented as a number). This actual ease is also associated with feelings of ease, comfort, and confidence in using numeric information. We suspect that both actual and felt ease influence quantification fixation, but leave it to future work to disentangle the two.

An interesting open question is when quantification fixation may be eliminated or reversed. There may be contexts in which people willfully choose to underweight numeric information, for instance. In domains where assigning numeric values to attributes feels distasteful (e.g., when making moral or ethical tradeoffs or when deciding between romantic partners), evaluators may find it aversive to rely on quantitative information and may overrely on available qualitative information instead.

A key implication of our findings is that when making decisions, people are systematically biased to favor options that dominate on quantified dimensions. And tradeoffs that pit quantitative against qualitative information are everywhere. Websites facilitating comparisons of options present us with a mix of quantified and nonquantified attributes to consider (e.g., price, star ratings). What's more, when making important decisions ranging from which medical treatment to use to whom we will hire, some attributes are more often quantified than others. When deciding between cancer treatments, people may face tradeoffs between their expected longevity and quality of life, only one of which is naturally represented as a number. In the workplace, when weighing diversity and inclusion priorities, an organization's diversity is much easier to quantify than its inclusion. Similarly, salary and paid time off are easily presented as numbers, while a company's culture is harder to quantify. Those who structure decision contexts ignore quantification fixation at their peril. As quantification becomes increasingly prevalent, people may be pulled away from valuable qualitative information toward potentially less diagnostic numeric information.

Methods

All experiments reported in this paper were approved by the University of Pennsylvania's Institutional Review Board (ID: 849979) or the Booth School of Business at the University of Chicago's Institutional Review Board (ID: 23-1348) and comply with all relevant ethical regulations. Informed consent was obtained from participants in all experiments reported (except where we obtained a waiver of informed consent for *SI Appendix, Experiment S7*), and each experiment had a distinct participant sample. The experimental data analyzed in this paper were collected via Qualtrics surveys (and condition assignment was random in all experiments, with randomization administered by Qualtrics) using Amazon Mechanical Turk (*Experiments 1a and 4* and *SI Appendix, Experiment S5b*), Prolific (*Experiments 1b, 1c, 2, and 3a* and *SI Appendix, Experiments S1, S2, S3a, S3b, S3c, S4, S5a, S8, S9a, and S9b*), in-person samples (*Experiment 3b*), Qualtrics Panels for a nationally representative sample of adults in the United States (*Experiment 5*), and Meta's Ad Manager (*SI Appendix, Experiment S7*). All experiments were preregistered.

Experiment 1a.

Participants. We recruited 1,000 participants[‡‡] on Amazon Mechanical Turk and paid them \$0.32 to complete a 2-min survey (434 women, 557 men, 5 nonbinary, 3 another gender, 1 prefer not to say; M_{age} = 42.71 y, SD = 12.03; 725 self-identified as monoracial white).
Analysis. Following our preregistered analysis plan, we ran a two-sample, two-tailed proportions test comparing how many participants chose the higher rated but more expensive hotel (Hotel Luxe) across conditions.

[‡‡]See *SI Appendix* for details about attrition rates for all experiments.

Experiment 1b.

Participants. We recruited 1,000 participants on Prolific and paid them \$0.32 to complete a 2-min survey (524 women, 461 men, 14 nonbinary, 1 another gender; M_{age} = 40.84 y, SD = 13.20; 676 self-identified as monoracial white).
Analysis. Following our preregistered analysis plan, we ran a two-sample, two-tailed proportions test comparing how many participants chose the candidate with the higher management grade but lower calculus grade across conditions.

Experiment 1c.

Participants. We recruited 1,000 participants on Prolific and paid them \$0.32 to complete a 2-min survey (486 women, 496 men, 17 nonbinary, 1 another gender; M_{age} = 36.89 y, SD = 12.84; 694 self-identified as monoracial white).
Analysis. Following our preregistered analysis plan, we ran a two-sample, two-tailed proportions test comparing how many participants chose the better connected but less sustainable conference location across conditions.

Experiment 2.

Participants. We recruited 2,000 participants on Prolific and paid them \$0.32 to complete a 2-min survey (964 women, 1,006 men, 29 nonbinary, 1 another gender; M_{age} = 41.05 y, SD = 13.13; 1,337 self-identified as monoracial white).
Analysis. Following our preregistered analysis plan, we ran an ordinary least squares (OLS) regression with robust SE to predict whether participants chose the employee with the higher likelihood of advancement but lower likelihood of retention. Our primary predictors were indicators for assignment to the *advancement quantified* condition, *both quantified* condition, and *neither quantified* condition (the *retention quantified* condition was the omitted comparison group). We used Wald tests to recover comparisons of interest across conditions.

Experiment 3a.

Participants. We recruited 1,000 participants on Prolific and paid them \$0.40 to complete a 3-min survey (490 women, 494 men, 13 nonbinary, 3 another gender; M_{age} = 36.89 y, SD = 13.25; 681 self-identified as monoracial white). Each participant earned a bonus based on their decision in the experiment (bonuses ranged from \$0.10 to \$0.40).
Methods. The candidates that were evaluated in this experiment were all real Prolific workers who had previously completed three different, 10-question games: (1) the Math Game [which involved a mental rotation task; (33)], (2) the Angles Game [which involved judging the angle and position of lines; (34)], and the (3) Trivia Game (which involved multiple-choice questions on various topics). These candidates were paid for each question they answered correctly, so they were motivated to perform well on each game.
Analysis. Following our preregistered analysis plan, we ran a two-sample, two-tailed proportions test comparing how many participants chose the candidate with the higher Math score (but lower Angles score) across conditions.

Experiment 3b.

Participants. We recruited participants across three different sites: two university campus labs (the Wharton Behavioral Lab = 253, University of Chicago Campus Lab = 82) and a local Chicago pop-up lab/storefront (Mindworks = 366). At both campus labs, participants were paid \$1.00 to complete a 3-min survey. At the local pop-up lab/storefront, participants were given 100 points (roughly equivalent to \$1.00) to exchange for prizes for completing our survey. Our experiment ran for three weeks (from April 26, 2024 until May 17, 2024). Our final sample consisted of 701 participants (445 women, 237 men, 10 nonbinary, 9 another gender; M_{age} = 28.28 y, SD = 12.43; 238 self-identified as monoracial white).
Analysis. Following our preregistered analysis plan, we ran an OLS regression with robust SE predicting whether participants chose the charity with the higher Accountability and Finance score but lower Culture and Community score (i.e., The Natural Resources Defense Fund) with an indicator for assignment to the *accountability and finance quantified* condition. Our regression included fixed effects for the site where data collection took place.

Experiment 4.

Participants. We recruited 2,000 participants on Amazon Mechanical Turk and paid them \$0.32 to complete a 2-min survey (1,044 women, 932 men, 23 nonbinary, 1

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.

another gender; M_{age} = 42.23 y, SD = 12.34; 1,463 self-identified as monoracial white).

Analysis. Following our preregistered analysis plan, we ran an OLS regression with robust SE to predict whether participants chose the higher benefit but less efficient project proposal. Our primary predictors were an indicator for assignment to the *benefit quantified* conditions, an indicator for assignment to the *disfluent number* conditions, and the interaction between these two indicators.

Experiment 5.

Participants. We recruited a sample of 602 adults through Qualtrics Panels who were representative of the United States population in terms of their age (M_{age} = 51.39, SD = 18.69), gender (56.8% women, 42.2% men, 1% nonbinary/another gender), race (65.8% self-identified as monoracial white), home geographic region (19.3% Northeast, 20.9% Midwest, 20.6% West, 39.2% South), and education (61.8% without a college degree) in September and October of 2023 to complete a 5-min survey.[§§] Each participant was truthfully informed that their decisions in our survey would determine which of two charities would receive a $1.00 donation.

Numeracy Measures. To capture objective numeracy (M = 1.25, SD = 0.94), following ref. 27 we asked participants to respond to a 4-item numeric understanding measure (4-NUM; (27)).[¶¶] Items were scored 1 if they were answered correctly, 0 if they were answered incorrectly, and summed to create a composite score. To assess subjective numeracy, following ref. 28 we asked participant to respond to an eight-item subjective numeracy scale [Cronbach's $\alpha = 0.86$; (28)] that included two subscales: the ability subscale (Cronbach's $\alpha = 0.91$; example item: "How good are you at calculating a 15% tip?") and the preference subscale (Cronbach's $\alpha = 0.71$; example item: "How often do you find numerical information to be useful?") All scale items used are referenced in *SI Appendix*. Participants responded to all subjective numeracy items on six-point Likert scales. Items were reverse-coded if necessary and then, following past research (28), averaged to create a composite. Participants responded to the objective numeracy and subjective numeracy scales in randomized order, and within the subjective numeracy scale, each of the subscales also appeared in randomized order.

Analysis. Following our preregistered analysis plan, we ran a two-sample two-tailed proportions test comparing how many participants chose to donate to the charity with the higher accountability and finance score but lower culture and community score (i.e., The Natural Resources Defense Fund) across conditions.

Following our preregistered exploratory analysis plan, we also ran three separate OLS regressions with robust SE (shown in Table 2) to predict whether each participant chose to donate to The Natural Resources Defense Fund as a function of their experimental condition and explored whether quantification fixation was moderated by participants' measured numeracy (either objective or subjective).

Data, Materials, and Software Availability. Fully anonymized data, materials, preregistrations, and analysis code are available on the Open Science Framework (https://osf.io/97peh/?view_only=566b843cf41a46f9962237423078597e) (35). All other data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. This research was supported by a MindCORE Postdoctoral Research Fellowship awarded to L.W.C. We are grateful to the Wharton Behavioral Lab, University of Chicago Campus Lab, and Mindworks for their help with data collection for *Experiment 3b*. We are grateful to members of the Duckworth-Milkman Lab at the University of Pennsylvania, members of the Behavior Change for Good Initiative at the University of Pennsylvania, members of the Honesty, Opportunity, Prosociality, & Ethics Lab at the University of Chicago, attendees at the Society for Judgment and Decision-making conference, attendees at the University of Pennsylvania's MindCORE seminar talk series, and attendees at the Booth Behavioral Faculty Brown Bag for their feedback. We are also grateful to J. Voelkel, D. Soman, and members of D. Soman's Psychology of Judgment and Decision-making class (G. Agarwal, C. Asuzu, A. Bambardekar, S. Merchant, R. Salamat, D. Turetski, and M. Yang) for their feedback on earlier versions of this manuscript. We are especially grateful to A. Smith for his invaluable research assistance.

[§§]See *SI Appendix* for full details of the demographic composition of participants. We requested a nationally representative sample of 600 US adults from Qualtrics Panels. Ultimately, our data included 602 adults instead. Our results remain consistent in direction, significance, and effect size whether or not we include the last two adults in our sample.

[¶¶]To protect the scale's usefulness to researchers, the authors have requested that researchers do not post any of the items online or in publications. Please see their original publication for more information.

Author affiliations: [a]Operations, Information, and Decisions Department, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104; [b]Behavioral Science Department, Booth School of Business, University of Chicago, Chicago, IL 60637; [c]Department of Economics, Massachusetts Institute of Technology, Cambridge, MA 02142; and [d]Electrical Engineering & Computer Science Department, Massachusetts Institute of Technology, Cambridge, MA 02139

1. Organisation of Economic Cooperation and Development, *OECD skills outlook 2013: First results from the survey of adult skills* (OECD Publishing, Paris, France, 2013).
2. S. Mamedova, E. Pawlowski, "Adult numeracy in the United States" (NCES 2020-025, National Center for Educational Statistics, 2020).
3. T. H. Freling, Z. Yang, R. Saini, O. S. Itani, R. R. Abualsamh, When poignant stories outweigh cold hard facts: A meta-analysis of the anecdotal bias. *Organ. Behav. Hum. Decis. Processes* **160**, 51–67 (2020).
4. D. A. Small, G. Loewenstein, P. Slovic, Sympathy and callousness: The impact of deliberative thought on donations to identifiable and statistical victims. *Organ. Behav. Hum. Decis. Processes* **102**, 143–153 (2007).
5. E. Peters *et al.*, Despite high objective numeracy, lower numeric confidence relates to worse financial and medical outcomes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 19386–19391 (2019).
6. E. Peters *et al.*, Numeracy and decision making. *Psychol. Sci.* **17**, 407–413 (2006).
7. M. Lipkus, E. Peters, G. Kimmick, V. Liotcheva, P. Marcom, Breast cancer patients' treatment expectations after exposure to the decision aid program adjuvant online: The influence of numeracy. *Med. Decis. Making* **30**, 464–473 (2010).
8. E. Peters, B. Shoots-Reinhard, "Better decision making through objective numeracy and numeric self-efficacy" in *Advances in Experimental Social Psychology*, B. Gawronski, Ed. (Academic Press, 2023), **vol. 68**, pp. 1–75.
9. A. Tversky, A. Intransitivity of preferences. *Psychol. Rev.* **76**, 31 (1969).
10. M. A. Lawson, R. P. Larrick, J. B. Scoll, When and why people perform mindless math. *Judgment Decis. Making* **17**, 1208–1228 (2022).
11. K. Hsee, Y. Yang, Y. Gu, J. Chen, Specification seeking: How product specifications influence consumer preference. *J. Consum. Res.* **35**, 952–966 (2009).
12. A. X. Yang, C. K. Hsee, Y. Liu, L. Zhang, The supremacy of singular subjectivity: Improving decision quality by removing objective specifications and direct comparisons. *J. Consum. Psychol.* **21**, 393–404 (2011).
13. L. Alter, D. M. Oppenheimer, Uniting the tribes of fluency to form a metacognitive nation. *Pers. Soc. Psychol. Rev.* **13**, 219–235 (2009).
14. D. M. Oppenheimer, The secret life of fluency. *Trends Cogn. Sci.* **12**, 237–241 (2008).
15. K. Shah, D. M. Oppenheimer, Easy does it: The role of fluency in cue weighting. *Judgment Decis. Making* **2**, 371–379 (2007).
16. L. K. Graf, S. Mayer, J. R. Landwehr, Measuring processing fluency: One versus five items. *J. Consum. Psychol.* **28**, 393–411 (2018).
17. C. K. Hsee, G. F. Loewenstein, S. Blount, M. H. Bazerman, Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychol. Bull.* **125**, 576 (1999).
18. C. K. Hsee, Y. Rottenstreich, Z. Xiao, When is more better? On the relationship between magnitude and subjective value *Curr. Dir. Psychol. Sci.* **14**, 234–237 (2005).
19. C. K. Hsee, The evaluability hypothesis: An explanation for preference reversals between joint and separate evaluations of alternatives. *Organ. Behav. Hum. Decis. Processes* **67**, 247–257 (1996).
20. K. Hsee, J. Zhang, General evaluability theory. *Perspect. Psychol. Sci.* **5**, 343–355 (2010).
21. J. M. Krijnen, D. Tannenbaum, C. R. Fox, Choice architecture 2.0: Behavioral policy as an implicit social interaction. *Behav. Sci. Policy* **3**, 1–18 (2017).
22. R. McKenzie, J. D. Nelson, What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychol. Bull. Rev.* **10**, 596–602 (2003).
23. S. Sher, C. R. McKenzie, Information leakage from logically equivalent frames. *Cognition* **101**, 467–494 (2006).
24. D. Tannenbaum, C. J. Valasek, Incentivizing wellness in the workplace: Sticks (not carrots) send stigmatizing signals. *Psychol. Sci.* **24**, 1512–1522 (2013).
25. K. A. Burson, R. P. Larrick, J. G. Lynch Jr., Six of one, half dozen of the other: Expanding and contracting numerical dimensions produces preference reversals. *Psychol. Sci.* **20**, 1074–1078 (2009).
26. W. Fagen-Ulmschneider, Perception of Probability Words (Version 1, GitHub, 2019). https://github.com/wadefagen/datasets/tree/master/Perception-of-Probability-Words. Accessed 5 February 2023.
27. M. Silverstein, P. Bjälkebring, B. Shoots-Reinhard, E. Peters, The numeric understanding measures: Developing and validating adaptive and nonadaptive numeracy scales. *Judgment Decis. Making* **18**, E19 (2023).
28. A. Fagerlin *et al.*, Measuring numeracy without a math test: Development of the subjective numeracy scale. *Med. Decis. Making* **27**, 672–680 (2007).
29. V. F. Reyna, C. J. Brainerd, Numeracy, gist, literal thinking and the value of nothing in decision making. *Nat. Rev. Psychol.* **2**, 421–439 (2023).
30. C. Camerer, M. Weber, Recent developments in modeling preferences: Uncertainty and ambiguity. *J. Risk Uncertainty* **5**, 325–370 (1992).
31. C. R. Fox, A. Tversky, Ambiguity aversion and comparative ignorance. *Q. J. Econ.* **110**, 583–603 (1995).
32. E. Peters, D. M. Markowitz, A. Nadratowski, B. Shoots-Reinhard, Numeric social-media posts engage people with climate science. *PNAS Nexus* **3**, pgae250 (2024).
33. G. Ganis, R. Kievit, A new set of three-dimensional shapes for investigating mental rotation processes: Validation data and stimulus set. *J. Open Psychol. Data* **3**, e3 (2015).
34. M. L. Collaer, S. Reimers, J. T. Manning, Visuospatial performance on an internet line judgment task and potential hormonal markers: Sex, sexual orientation, and 2D: 4D. *Arch. Sexual Behav.* **36**, 177–192 (2007).
35. L. W. Chang, E. L. Kirgios, S. Mullainathan, K. L. Milkman, Data from "Does counting change what counts? Quantification fixation biases decision-making." Open Science Framework. https://osf.io/97peh/?view_only=566b843cf41a46f9962237423078597e. Deposited 3 September 2024.

Downloaded from https://www.pnas.org by MASSACHUSETTS INSTITUTE OF TECHNOLOGY; MIT LIBRARIES on July 23, 2025 from IP address 18.11.82.240.