Boon Han Koh

Research Statement

October 2026

Senior Lecturer
Department of Economics
University of Exeter Business School

Email: b.koh@exeter.ac.uk
Web: http://www.boonhankoh.net

Overview

My fields of research are experimental and behavioral economics, and my research uses a combination of laboratory, online, and field experiments. A unifying question in my work is how people form beliefs about themselves and others, and how these beliefs influence behavior. I study (i) how these beliefs are biased by gender and identity; (ii) how the information people seek and receive shapes their beliefs and behavior; and (iii) how researchers can measure beliefs and identity reliably and sensitively. Some of my experiments are motivated by empirical patterns observed in practice, while others are designed to test theoretical predictions.

My research is highly collaborative. I have built an extensive network of collaborators from different regions, including senior co-authors and my peers, and work closely with them on many projects. I have also worked on many joint projects with PhD students and early-career researchers, mentoring them through collaboration.

1 Discrimination and Identity

1.1 Gender, identity, and biases in evaluations of self and others

In many environments, the outcomes of decision makers depend on both their actions, which are often unobserved by others, and luck. This uncertainty creates scope for evaluators to form biased beliefs, including those that depend on the decision maker's gender. Such biases may contribute to the observed gender gaps in labor market outcomes. My research examines these biases in a context where decision makers, such as leaders, make choices that affect the outcomes of others in their group. How are leaders' outcomes evaluated in such environments, and are these evaluations biased by the leader's gender? The first question is the basis of the work I did during my PhD: In "**By chance or by choice? Biased attribution of others' outcomes when social preferences matter**" (Erkal, Gangadharan, and Koh, 2022, **Experimental Economics**), we find that evaluators attribute positive outcomes more to luck and negative outcomes more to the leader's actions.

This project then led to a successful **Australian Research Council** grant bid post-PhD (Discovery Project, **276,611 AUD**), through which I developed a research agenda with my co-authors to investigate whether the gender of the leader also influences the evaluation process. Several projects have grown out of this agenda. In "**Do women receive less blame than men? Attribution of outcomes in a prosocial setting**" (Erkal, Gangadharan, and Koh, 2023, **Journal of Economic Behavior & Organization**), we ask whether evaluators attribute the outcomes of male and female leaders differently. We document gender biases in the attribution of negative outcomes: Negative outcomes of male (female) leaders are attributed more to their selfish decisions (bad luck), a pattern driven mainly by male evaluators. Beyond documenting these gender biases in beliefs and attribution of outcomes, we also ask whether and how these beliefs translate into the bonuses and penalties received by leaders. In "**Gender criteria gap in evaluation: role of perceived intentions and outcomes**" (Erkal, Gangadharan, and Koh, 2026a, **International Economic Review**), we make a novel contribution

by showing that a **gender criteria gap** exists in evaluations. When evaluators are able to make discretionary payments after they have formed beliefs about the leaders' intentions, the payments made to male leaders are determined by both outcomes and perceived intentions, but those made to female leaders are determined by outcomes only. This implies that men and women are subject to different criteria in the evaluation process. Taken together, our findings suggest that gender gaps can arise not only from biased beliefs about leaders' actions, but also from differences in how evaluators use these beliefs.

My future work in this area will investigate (i) whether evaluators choose to learn about different criteria (e.g., ability versus effort) for men and women when assessing candidates in a hiring context, and whether any such gender differences lead to suboptimal hiring decisions; and (ii) whether asking evaluators to privately reflect on or publicly justify their compensation decisions mitigates gender biases in the evaluation process by increasing accountability. We are currently developing research designs to address both sets of questions.

This research agenda has also led to an **interdisciplinary collaboration with medical doctors** at Norfolk and Norwich University Hospital (NNUH), where we investigate gender biases in the evaluation of trainees in anesthesia, and whether speaking with confidence benefits trainees differently by gender. Using a series of vignette studies with consultants in anesthesia, our preliminary analysis suggests that there exists a confidence premium, such that trainees who come across as more confident receive more favorable evaluations. However, confidence also affects the evaluation of outcomes differently for men and women. Confident women are penalized less for failures than unconfident women, while confident men are rewarded more for successes than unconfident men. Hence, confidence benefits women when they fail and men when they succeed (Erkal et al., WIPb). We are currently running a survey with trainees to examine whether they anticipate these biases in evaluations, and whether these beliefs translate into how much confidence they project. We expect this research to have direct policy implications for how medical doctors are evaluated and trained in their profession.

Apart from understanding how individuals are evaluated by others, another important question is whether men and women evaluate themselves differently. Gender biases in self-evaluations can affect individuals' confidence, which may in turn influence the career choices they make, thus contributing to occupational segregation and gender gaps in leadership. The gender gap in self-evaluation could simply reflect men and women having different preferences across task domains. In "**Does domain preference explain gender gaps in self-evaluation?**" (Aksoy, Chadd, and Koh, 2026) (revise-and-resubmit at **European Economic Review**), we study whether the gender gaps in self-evaluation in male-stereotyped domains persist when accounting for domain preference. We replicate the findings of Exley and Kessler (2022) that women have lower confidence and provide harsher self-evaluations than equally performing men. Examining individuals' domain preferences, we find that men have a stronger preference for math and science domains than women, even after controlling for ability. However, controlling for this domain preference gap does not reduce the gender gaps in self-evaluation. Our findings suggest that policymakers need to focus on the root causes of the persistent gender gap in self-evaluation, which may include gender stereotypes, social norms, or experiences of discrimination, among others.

Under this research agenda, I also have two ongoing projects with early-career researchers on gender differences in perceptions of persistence and in credit claiming. Gender stereotypes can contribute to incorrect beliefs, which may in turn lead to unfair treatment of men and women. In the first project, joint with Emily Tran (PhD student at the University of Melbourne), we ask whether individuals hold correct perceptions about the persistence of men and women. We show that individuals hold misperceptions about men's and women's persistence following positive and negative performance feedback in male- and female-stereotyped tasks (Erkal et al., WIPc). Beliefs may matter for how credit is shared in teams, since credit claiming depends on

what beliefs individuals hold about their own contributions and those of others. In the second project, joint with Xiaojie Zhang (PhD student at Nanyang Technological University), we ask whether men and women claim credit differently in team tasks when teams differ in their level of specialization, which may affect how easy or difficult it is for contributions to be attributed to individual team members. We find that, while women tend to claim less credit than men on average, this gender gap in credit claiming does not depend on specialization (Koh, Yeo, and Zhang, WIP). Data collection has been completed for both projects, and we are currently preparing the manuscripts for submission.

In more recent work, I am interested in how the design of institutions can contribute to biased evaluations. In the past decade, there has been a greater push toward diversity, equity, and inclusion (DEI) policies in organizations, such as quotas or rules that favor minority candidates in hiring. What is less understood is how evaluators respond to these policies. Together with Romain Gauriot and Francisco Silva, I examine this question both theoretically and experimentally. Our preliminary theoretical analysis suggests that, under certain conditions, a tie-break rule that favors minority candidates may distort evaluators' assessments and backfire, reducing the likelihood that minority candidates are hired. We are currently developing an experimental framework to test these theoretical predictions. We have submitted a grant bid to the **Australian Research Council** (**Discovery Project**, outcome pending) to further develop this work.

My work and expertise in this research agenda have also led to three encyclopedia entries on gender biases in beliefs and attribution of outcomes (Erkal and Koh, 2025; Erkal, Gangadharan, and Koh, 2026c,b). I have also **guest-edited a special issue on Discrimination and Diversity** in the **Journal of Economic Behavior & Organization** (Borcan et al., 2025).

1.2 Sexual and gender minorities

Another strand of my research on identity and discrimination focuses on discrimination against and attitudes toward gender and sexual minorities. Discrimination against LGBTQ+ people is hard to measure because sexual identity is usually invisible and often not disclosed. As a result, we know little about how anticipated discrimination shapes minorities' behavior, such as whether they disclose their identity. In "**Sexual identity, gender, and anticipated discrimination in prosocial behavior**" (Aksoy, Chadd, and Koh, 2023, **European Economic Review**), we study whether individuals strategically mask signals about their affinity with the LGBTQ+ community in response to anticipated discrimination in prosocial behavior. In an online dictator game, recipients are given the opportunity to signal their LGBTQ+ affinity either before or after the dictators make their allocation decisions. A key innovation in this paper is the novel **Icon Task** that allows individuals to signal their LGBTQ+ affinity in a non-intrusive manner. We find that women are more likely to hide such signals when they learn that dictators are able to see these signals prior to making decisions. Analysis using data from the American National Election Studies suggests that this gender difference is likely due to women's higher propensity to anticipate discrimination, implying that women may hold more pessimistic beliefs about the prevalence of discrimination. However, we find that dictators do not treat recipients differently based on signals of their LGBTQ+ affinity, suggesting that such strategic masking of identity may be unnecessary. Our findings have been featured in WAMC's 51 Percent Podcast.

Building on my expertise in experimental economics and in LGBTQ+ research, I co-authored a handbook chapter that is the first of its kind to review the experimental economics literature on gender and sexual minorities (Aksoy et al., 2027). Most existing evidence on the experiences of LGBTQ+ individuals comes from Western, high-income countries, with much less known about non-WEIRD contexts. To address this gap, I have led an **interdisciplinary project**, funded by the **British Council** (Gender Equality Partnerships, **24,852 GBP**; as PI), that combines

qualitative and quantitative data to examine the lived experiences of LGBTQ+ individuals in Vietnam (Evangelista et al., 2025).

2 Information Demand and Provision

A second major theme of my research examines how people seek, provide, and respond to information about performance, and how this shapes the beliefs and behavior of individuals and teams. Standard economic models predict that individuals value information because it helps them make better decisions. In practice, however, people often avoid useful information or withhold it from others. Such behavior may arise from motivated belief updating and the desire to protect one's ego (Golman, Hagmann, and Loewenstein, 2017; Golman et al., 2022). Why do individuals avoid information or withhold it from others, and how does this depend on whether the information is instrumental for decision making? A key feature of my research is that it experimentally varies the instrumentality of information, which allows us to separate its instrumental value for decision making from the psychological motives driving information provision or demand.

On the supply of information, biases in feedback provision in the workplace can have a detrimental impact on career progression, as the recipients of feedback may lack the information needed to improve or to make important career decisions. Such biases can also exacerbate gender gaps in labor market outcomes if feedback is provided differently to men and women in the workplace. In "**The signals we give: Performance feedback, gender, and competition**" (Coutts, Koh, and Murad, 2026, **Management Science**), we study feedback provision across variations in the nature of performance signals, their instrumental value, and the recipient's gender. We find that a surprising amount of both positive and negative feedback is withheld. Using large language models (LLMs) to conduct text analysis of survey responses in a follow-up experiment, we find that advisors' feedback decisions are driven mainly by transparency- and duty-related considerations, and, to a lesser extent, by self-serving motives. Moreover, we find that advisors are more likely to withhold vague, non-instrumental negative feedback than positive feedback, with suggestive evidence that this is more pronounced for female recipients due to stronger ego-protective concern by advisors toward women than men. This project was funded by the **British Academy** (Leverhulme Small Research Grants, **9,974 GBP**; as PI).

On the demand side, understanding individuals' demand for information is important for the effective design of feedback systems, especially if individuals seek or avoid information when it is optimal for them to do otherwise. I have two current projects that examine this question. The first is a project with my former PhD student (Lingguo Xu, now a postdoctoral research fellow at the University of Waterloo) that examines gender differences in the demand for social comparison information. In "**Gender and demand for social comparison information**" (Erkal, Koh, and Xu, 2026) (revise-and-resubmit at **Journal of Economic Psychology**), we show that men and women are no different in their demand for social comparison information both when this information is potentially instrumental and when it is non-instrumental. A second project, in which we have just finished data collection, examines individuals' demand for feedback about their own performance depending on both the valence of performance signals and the instrumentality of feedback (Erkal et al., WIPa). Our preliminary analysis shows that individuals who receive positive signals of performance are less likely to seek feedback when it is instrumental than when it is non-instrumental. Moreover, when feedback is instrumental, providing information can be welfare-improving for those who do not demand such information, suggesting that institutions may sometimes need to override individuals' preferences to improve their outcomes.

The same tension between individual preferences and welfare-improving information provision may also arise in teams. In "**Milestones work, but teams don't choose them:**

Information provision and self-governance in teams" (Erkal, Koh, and Lam, 2026), joint with Nguyen Lam, who began the project as an undergraduate student (now a PhD student at Northwestern University), we study how institutions that govern information affect cooperation in teams, and whether these institutions should be externally imposed (external governance) or chosen by team members themselves (self-governance). Across three sets of laboratory experiments involving 1,077 participants, we find that providing teams with updates using predetermined milestones under external governance significantly improves cooperation. However, when given a choice, only a quarter of members vote for milestone-based feedback. This feedback structure is disproportionately favored by self-interested individuals. Analysis of participants' survey responses using LLMs suggests that these members want to conceal their free-riding behavior, while conditional cooperators prefer regular updates so that they can monitor their teammates and adjust their contributions accordingly. Consequently, teams perform worse under self-governance, with milestone-based feedback leading to the worst outcomes for teams when endogenously chosen.

Our findings have direct implications for how organizations should design progress updates in teams. The project was funded by the **British Academy** (Leverhulme Small Research Grants, **9,905 GBP**; as PI), and a broader research agenda is funded by the **Australian Research Council** (Discovery Project, **436,894 AUD**). **My future work under this research agenda** will investigate (i) whether leaders can implement welfare-improving feedback provision institutions for teams; (ii) the effectiveness of different feedback and governance structures in promoting cooperation within diverse teams; and (iii) the role of information interventions in achieving more effective cooperation under self-governance.

3 Experimental Methodology

A third theme of my research focuses on developing and evaluating experimental tools, both for studying vulnerable individuals and for eliciting beliefs. In Aksoy, Chadd, and Koh (2023), discussed above, we introduce the **Icon Task**, which lets participants anonymously signal their affinity with the LGBTQ+ community. This allows researchers to study gender and sexual minorities, a highly vulnerable group in many regions, in controlled experiments without requiring explicit disclosure. Building on this, our handbook chapter (Aksoy et al., 2027) provides a set of practical recommendations for collecting sexual orientation and gender identity data and conducting studies with LGBTQ+ individuals.

Because a large part of my research relies on eliciting beliefs from participants, I also study how the elicitation method can influence the reliability of the data collected. The binarized scoring rule (BSR) has become a standard tool for belief elicitation because it is theoretically robust to risk attitudes, but whether it elicits truthful reports in practice depends on assumptions that are rarely tested. I have one publication on this topic (Erkal, Gangadharan, and Koh, 2020, **Journal of Economic Psychology**), as well as a working paper and a work in progress that investigate distortions in belief reports under the BSR. In "**High stakes, more mistakes? Belief elicitation and incentives**" (Coutts et al., 2025), we focus on the role of financial incentives and show that errors in belief reports increase by 33% when financial incentives under the BSR are increased tenfold, suggesting that using lower incentives is preferable in belief-elicitation procedures. In "**Hedging ambiguity in belief elicitation**" (Coutts et al., WIP), we make a novel contribution by showing, both theoretically and experimentally, that the BSR is not robust to ambiguity attitudes. Specifically, we find that individuals with non-neutral ambiguity attitudes have an incentive to shift or distort their reports due to hedging motives. These findings have important implications for how researchers elicit participants' beliefs in surveys and experiments.

References

Aksoy, B., I. Chadd, B. Grosskopf, B. H. Koh (2027). Sexual orientation and gender identity. In C. Schwieren, M. Cubel (eds.), *Handbook of Gender and Experimental Economics*. Edward Elgar Publishing.

Aksoy, B., I. Chadd, B. H. Koh (2023). Sexual identity, gender, and anticipated discrimination in prosocial behavior. *European Economic Review*, 154:104427.

Aksoy, B., I. Chadd, B. H. Koh (2026). Does domain preference explain gender gaps in self-evaluation? *Working Paper*.

Borcan, O., L. Gee, L. Harvey, B. H. Koh, E. Reuben (2025). An introduction to the special issue on discrimination and diversity.

Coutts, A., C. Drobner, B. H. Koh, C. Woolnough (2025). High stakes, more mistakes? belief elicitation and incentives. *Working Paper*.

Coutts, A., C. Drobner, B. H. Koh, C. Woolnough (WIP). Hedging ambiguity in belief elicitation. *Working Paper*.

Coutts, A., B. H. Koh, Z. Murad (2026). The signals we give: Performance feedback, gender, and competition. *Management Science*.

Erkal, N., M. Fonseca, B. H. Koh, L. Lou (WIPa). Beyond the win: Luck and its influence on feedback seeking. *Work in Progress*.

Erkal, N., L. Gangadharan, S. Kausar, B. H. Koh, Z. Najim, M. Roe (WIPb). Speaking with confidence: Who benefits and when? evidence from physicians. *Work in Progress*.

Erkal, N., L. Gangadharan, B. H. Koh (2020). Replication: Belief elicitation with quadratic and binarized scoring rules. *Journal of Economic Psychology*, 81:102315.

Erkal, N., L. Gangadharan, B. H. Koh (2022). By chance or by choice? biased attribution of others' outcomes when social preferences matter. *Experimental Economics*, 25(2):413–443.

Erkal, N., L. Gangadharan, B. H. Koh (2023). Do women receive less blame than men? attribution of outcomes in a prosocial setting. *Journal of Economic Behavior & Organization*, 210:441–452.

Erkal, N., L. Gangadharan, B. H. Koh (2026a). Gender criteria gap in evaluation: role of perceived intentions and outcomes. *International Economic Review*.

Erkal, N., L. Gangadharan, B. H. Koh (2026b). Gender differences in attribution of others' outcomes. In B. Kebede (ed.), *Encyclopedia of Experimental Social Science*. Edward Elgar Publishing.

Erkal, N., L. Gangadharan, B. H. Koh (2026c). Gender differences in attribution of own outcomes. In B. Kebede (ed.), *Encyclopedia of Experimental Social Science*. Edward Elgar Publishing.

Erkal, N., L. Gangadharan, B. H. Koh, E. Tran (WIPc). Gender and perception of persistence. *Work in Progress*.

Erkal, N., B. H. Koh (2025). Gender and (inaccurate) beliefs. In S.-H. Chuah, R. Hoffmann, A. Neelim (eds.), *Edward Elgar Encyclopedia of Behavioral and Experimental Economics*. Edward Elgar Publishing.

Erkal, N., B. H. Koh, N. Lam (2026). Milestones work, but teams don't choose them: Information provision and self-governance in teams. *Working Paper*.

Erkal, N., B. H. Koh, L. Xu (2026). Gender and demand for social comparison information. *Working Paper*.

Evangelista, Z., F. Arico, B. H. Koh, P. Liamputtong, T. T. T. Lien (2025). LGBTQ+ student lives beyond weird contexts: Experiences in vietnamese higher education. *Working Paper*.

Exley, C. L., J. B. Kessler (2022). The gender gap in self-promotion. *The Quarterly Journal of Economics*, 137(3):1345–1381.

Golman, R., D. Hagmann, G. Loewenstein (2017). Information avoidance. *Journal of Economic Literature*, 55(1):96–135.

Golman, R., G. Loewenstein, A. Molnar, S. Saccardo (2022). The demand for, and avoidance of, information. *Management Science*, 68(9):6454–6476.

Koh, B. H., J. Yeo, X. Zhang (WIP). Gender, task specialization, and credit claiming. *Work in Progress*.