Blinding curiosity: Preferences for "blinding" one's own judgment

Sean Fath, Richard Larrick, & Jack Soll

Duke University

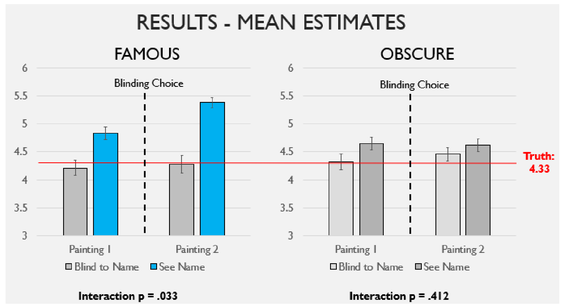
Overview & Hypotheses

- "Blinding" – purposefully restricting the information one incorporates into a decision in order to avoid being influenced adversely – can be useful, at a policy level, as a structural shield against bias (Goldin & Rouse, 2000).
- However, in the absence of blinding as policy, will people choose for themselves to be blind to potentially biasing information when making decisions, evaluations, etc? In the present research, we perform the first tests of individual-level blinding preferences.
- We hypothesize that people will be less inclined to blind themselves to potentially contaminative information when making blinding decisions *in the moment* than when considering the *appropriateness of blinding in a hypothetical sense*.
- We hypothesize that different prioritization of curiosity vs concerns over accuracy and fairness will explain a discrepancy between actual and normative blinding preferences.
- We test these hypotheses across 4 pre-registered studies (N = 1863).

Study 1a

Participants. 396 participants recruited from Amazon's Mechanical Turk.

Procedure. Study 1a explored actual blinding preferences for impending estimations, and demonstrated how viewing contaminative information can lead to bias. Participants guessed the average attractiveness rating given to two late-renaissance paintings in a pretest. Participants were informed that pretest participants rated the attractiveness of the paintings blind to artist identity. Participants guessed the average rating of the first painting blind to artist identity, but chose whether to learn the name of the second painting's artist before guessing its rating (blinding choice). Participants who chose to see the name were randomly provided with a famous (e.g., Michelangelo) or obscure (e.g., Ghirlandaio) artist's name. We expected that participants who chose to learn the name of the second painting's artist would provide upwardly biased estimates of its attractiveness, but only when provided with a famous name. The experiment had a 2 (Estimate: 1st painting vs. 2nd) x 2 (Fame: famous vs. obscure) x 2 (Choice: view name vs. blind) x 2 (Painting order) design. 60.4% of participants chose to learn the name of the second painting's artist. There was no effect of the order variable, which was collapsed. The predicted 3-way interaction emerged. $F(1, 392) = 4.57$, $p = .033$, $\eta p^2 = .012$. Follow-up analyses revealed participants gave more positively biased estimates of attractiveness ratings when they chose to see the name of the artist, but only when that name turned out to be famous.

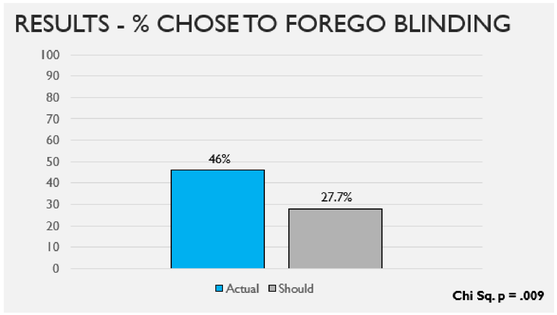


Study 1b

Participants. 204 participants were recruited from Mechanical Turk.

Procedure. Study 1b explored actual vs normative blinding preferences in the same context as Study 1a. Participants estimated the attractiveness of Painting 1 without being given a name; had the option to learn the name of the painter of Painting 2. Participants in the *actual* blinding preference condition responded to the question: "Would you like to know the name of the artist?" Participants in the *normative* blinding preference condition responded to the question: "Suppose your goal is to be as accurate as you

Ctd. can. Do you think you should know the name of the artist when you evaluate the next painting?" As predicted, more participants chose to learn the name in the *actual* vs *normative* ("should," below) condition.



Studies 1a & b demonstrated (i) more people choose to view contaminative information in the moment than when considering the appropriateness of viewing it in the abstract & (ii) choosing to view contaminative info can make decision-makers susceptible to bias and harm their accuracy.

Study 2a

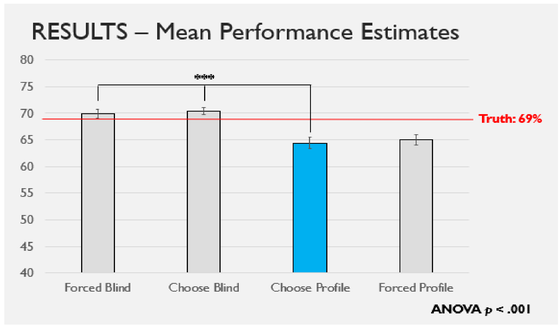
Participants. 804 participants were recruited from Prolific Academic.

Procedure. Study 2a explored actual blinding preferences and bias associated with viewing contaminative info, but in a performance evaluation context. Participants viewed a video of one of someone's performance in a pattern recognition task, and estimated the quality of the performance. Participants in the *forced blind* condition saw the video of the performance alone before evaluating the performance. Participants in the *forced profile* condition saw a profile of the performer before viewing the video (below), containing potentially contaminative information indicating a negative performance on a different, unrelated task. Participants in the *choice* condition chose whether to see the profile of the performer before watching the video, or be blind to the profile.

PROFILE

Name: Sean F
Hometown: Chicago, IL
Education: Community college graduate
College Major: English
Math Task Performance: 40% Correct

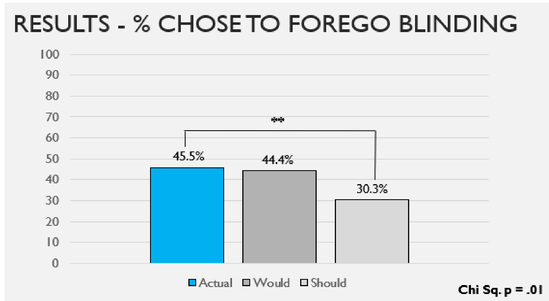
All participants guessed the quality of the performance on a 21-point scale (1 = 0%, 21 = 100%) and had the option to input more specific estimates or confirm their scale choice. We expected that those who chose to view the profile (or did so automatically) would report negatively biased estimates of the performance. In support of this, a 2 (Blinding Decision: ability to choose vs. automatic) x 2 (Blinding Outcome: see profile vs. blind) ANOVA test of bias revealed a main effect of Blinding Outcome, $F(1, 800) = 40.84$, $p < .001$, $\eta p2 = .05$, but not a main effect of Blinding Decision or an interaction (p's > .5)

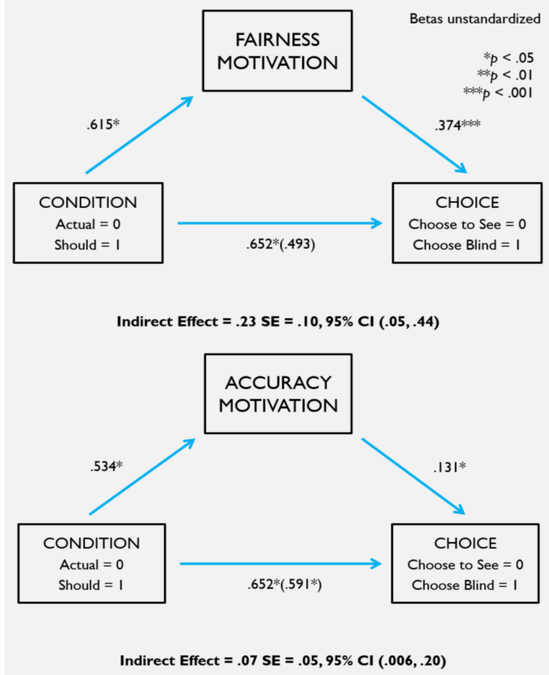


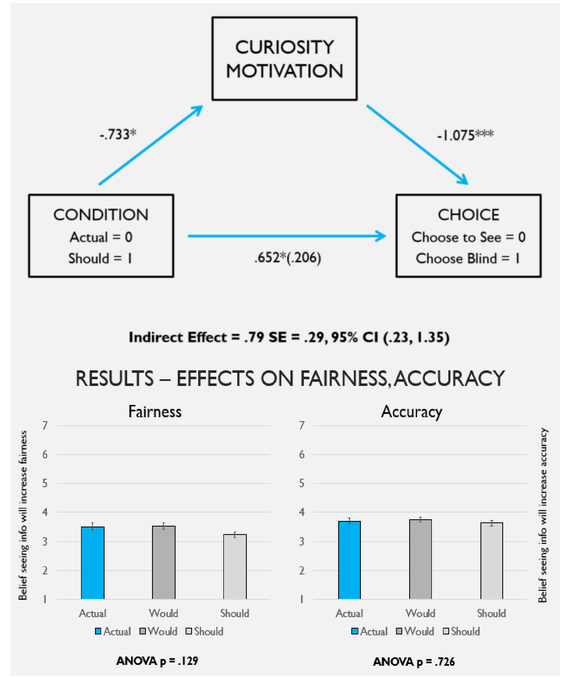
Study 2b

Participants. 459 participants were recruited from Prolific Academic.

Procedure. Study 2b compared actual, normative, and merely hypothetical blinding preferences in the same context as Study 2a, and also probed the motivations and lay beliefs underlying these preferences. Participants in the *actual* condition indicated their blinding choice (see profile or not) in the same manner as participants in the *choice* condition in Study 2a. Participants in the *should* (normative preference) condition were asked: "If you were in this scenario, would it be appropriate for you to view this person's profile information before you watch the video of their performance?" Participants in the *would* (hypothetical preference) were asked: "If you were in this scenario, would you want to view this person's profile information before you watch the video of their performance?" Three follow-up questions assessed the extent to which their blinding choice was motivated by concerns over fairness, concerns over accuracy, and curiosity. Two follow-up questions assessed participants naïve theories about the effect of foregoing blinding on the fairness and accuracy of their assessments. As in Study 1b, we expected more participants to choose to view the potentially contaminative info (the profile) when making an actual blinding choice than when making a normative blinding choice. We expected participants in the *actual* condition to be relatively more motivated by curiosity, and relatively less motivated by concerns over fairness and accuracy, than those in the *should* condition.







Summary

- Even though most people (~70%) indicate they *should* be blind when considering blinding decisions in a hypothetical sense, around 50% choose to view contaminative info when making a blinding decision in the moment.
- Choice to view info leaves decision-makers susceptible to bias and harms the accuracy of their judgments when they are biased.
- Those choosing in the moment are more motivated by curiosity, and less by fairness/accuracy concerns, than those making a normative choice. Curiosity predicts a choice to see contaminative info, whereas concerns over fairness/accuracy predict a choice to be blind
- People generally agree that seeing contaminative information will decrease the fairness and accuracy of their estimates. Regardless, some still choose to see it; curiosity gets the better of them.

CONTACT:

sean.fath@duke.edu