Bayesian Experimental Design with Mutual Information and Learned Errors for Human-Computer Interaction

Hugo Miquel[*,†], Julien Gori[‡], and Olivier Rioul[*]

[*]LTCI Télécom Paris, Institut Polytechnique de Paris, France
[†]ENS Paris-Saclay, Université Paris-Saclay, France
[‡]Sorbonne Université, CNRS, Inserm, ISIR, France

European Mathematical Psychology Group Meeting (EMPG 2025)

University of Padua, Italy
September 5th, 2025

Outline

Bayesian Information Gain (BIG) Framework

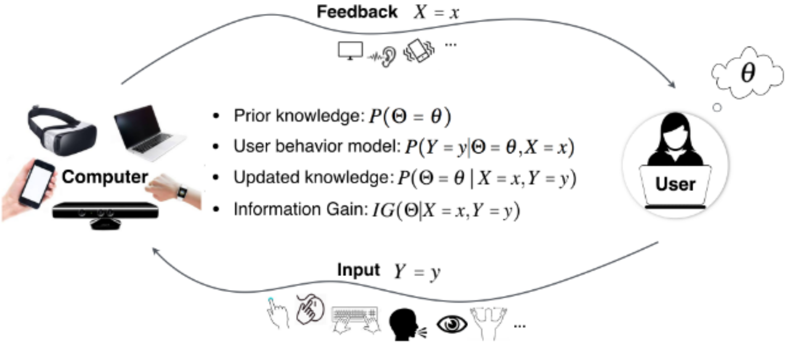
A special case of *Bayesian Experimental Design* [Liu+al CHI'2017]

Three key random variables:

- Θ: User's intended target
- X: System feedback
- Y: User input

Information gain:

- $IG(\Theta|X = x, Y = y) = \underbrace{H(\Theta)}_{\text{entropy}} - \underbrace{H(\Theta|X = x, Y = y)}_{\text{conditional entropy}}$



Bayesian Update in the BIG Framework

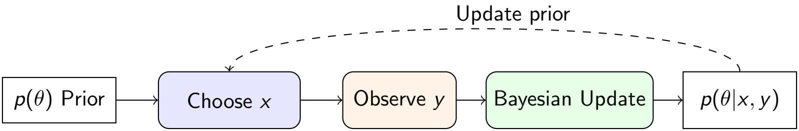


Figure 1: Bayesian update cycle: inference of θ with direct feedback

Posterior distribution:

$$p(\theta \mid x, y) = \frac{p(y \mid x, \theta) \cdot p(\theta)}{p(y \mid x)}$$

User behavior model (likelihood):

$$p(y \mid x, \theta)$$

The BIG Framework: Information-Theoretic Utility

Utility Function in BIG: Conditional Mutual Information

$$U(x) = I(\Theta; Y|X = x) = H(Y|X = x) - H(Y|\Theta, X = x)$$

Expected reduction in uncertainty about target Θ averaged over all possible responses Y

The BIG Framework: Information-Theoretic Utility

Utility Function in BIG: Conditional Mutual Information

$$U(x) = I(\Theta; Y|X = x) = H(Y|X = x) - H(Y|\Theta, X = x)$$

Expected reduction in uncertainty about target Θ averaged over all possible responses Y

Optimal Feedback Selection

$$x^* = \arg\max_x I(\Theta; Y|X = x)$$

Maximizing mutual information is a logical choice that likely reduces the expected number of interactions needed to identify the user's target.

Impact of User Errors on BIG

Zero Error Assumption

User model, likelihood:

$$p(y|x,\theta) = \begin{cases} 1, & \text{if } y = f(x,\theta) \\ 0, & \text{otherwise} \end{cases}$$

where $f(x,\theta)$ is the "correct" response

Impact of User Errors on BIG

Zero Error Assumption

User model, likelihood:

$$p(y|x,\theta) = \begin{cases} 1, & \text{if } y = f(x,\theta) \\ 0, & \text{otherwise} \end{cases}$$

where $f(x,\theta)$ is the "correct" response

If user makes an error by providing $y' \neq f(x,\theta^*)$:

$$p(\theta^*|x,y') = \frac{p(y'|x,\theta^*)p(\theta^*)}{p(y'|x)} = \frac{0 \cdot p(\theta^*)}{p(y'|x)} = 0$$

$\Rightarrow$ The true target θ^* is eliminated permanently!

Impact of User Errors on BIG

Zero Error Assumption

User model, likelihood:

$$p(y|x,\theta) = \begin{cases} 1, & \text{if } y = f(x,\theta) \\ 0, & \text{otherwise} \end{cases}$$

where $f(x,\theta)$ is the "correct" response

If user makes an error by providing $y' \neq f(x,\theta^*)$:

$$p(\theta^*|x,y') = \frac{p(y'|x,\theta^*)p(\theta^*)}{p(y'|x)} = \frac{0 \cdot p(\theta^*)}{p(y'|x)} = 0$$

$\Rightarrow$ The true target θ^* is eliminated permanently!

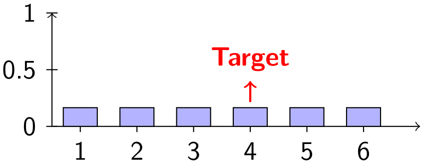
Need for Error-Robust Models

- With this user model, BIG is not resilient to user errors
- Need robust models that can recover from occasional errors

Example of User Error Impact

Binary Search Example: Target space $\Theta = \{1, 2, 3, 4, 5, 6\}$, true target $\theta^* = 4$

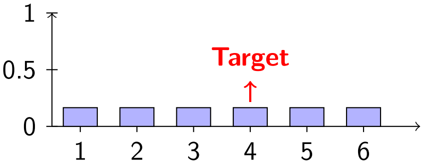
Initial distribution: Uniform prior
$p(\theta) = 1/6$



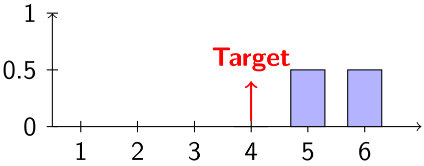
Example of User Error Impact

Binary Search Example: Target space $\Theta = \{1, 2, 3, 4, 5, 6\}$, true target $\theta^* = 4$

Initial distribution: Uniform prior $p(\theta) = 1/6$



After error: User asked "Is $\theta \leq 4$? " but incorrectly answers "No"



Consequence of Error with Zero Error Model:

- Update: $p(\theta) = 0$ for $\theta \in \{1, 2, 3, 4\}$, $p(\theta) = 1/2$ for $\theta \in \{5, 6\}$
- System cannot recover without starting over

Fixed Error rate Model

Error Model Parameters

- ϵ_0: Error rate parameter ($0 \leq \epsilon_0 \leq 1$)
- $q(y|x)$: Distribution of errors (often uniform on incorrect responses)

Likelihood Function with Error Parameter

$$p(y|x, \theta, \epsilon_0) = (1 - \epsilon_0) \cdot \delta(y, f(x, \theta)) + \epsilon_0 \cdot q(y|x)$$

Fixed Error rate Model

Error Model Parameters

- ϵ_0: Error rate parameter ($0 \leq \epsilon_0 \leq 1$)
- $q(y|x)$: Distribution of errors (often uniform on incorrect responses)

Likelihood Function with Error Parameter

$$p(y|x, \theta, \epsilon_0) = (1 - \epsilon_0) \cdot \delta(y, f(x, \theta)) + \epsilon_0 \cdot q(y|x)$$

Effect on Posterior Update

With $\epsilon_0 > 0$, even if $y' \neq f(x, \theta^*)$:

$$p(\theta^*|x, y') = \frac{\epsilon_0 \cdot q(y'|x) \cdot p(\theta^*)}{p(y'|x)} > 0$$

Key Insight: The user target probability decreases but remains non-zero! This enables recovery from user errors.

Parameter Mismatch Problem in BIG

Critical Challenge: What happens when the user's true error rate ϵ^* differs from our model assumption ϵ_0?

- **Case 1:** $\epsilon^* < \epsilon_0$ (Overestimation)
  - System becomes unnecessarily cautious
  - **Result:** High accuracy but excessive queries
  - System attributes less confidence to correct user responses

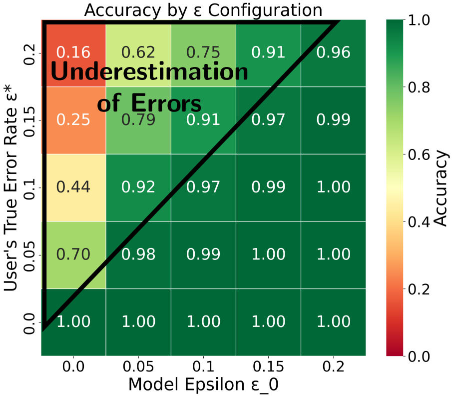
Parameter Mismatch Problem in BIG

Critical Challenge: What happens when the user's true error rate ϵ^* differs from our model assumption ϵ_0?

- **Case 1:** $\epsilon^* < \epsilon_0$ (Overestimation)
  - System becomes unnecessarily cautious
  - **Result:** High accuracy but excessive queries
  - System attributes less confidence to correct user responses
- **Case 2:** $\epsilon^* > \epsilon_0$ (Underestimation)
  - System trusts user responses too much
  - **Result:** Reduced accuracy, potential failure
  - Errors have stronger impact on posterior distribution

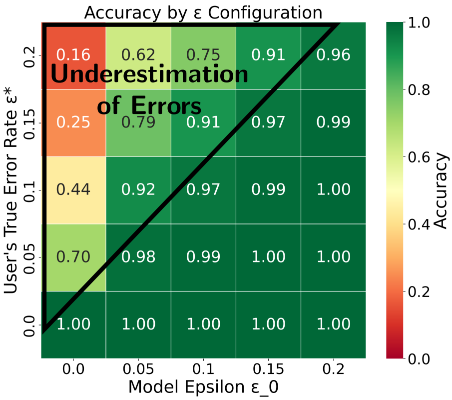
Key Problem: The fixed error rate model requires accurate knowledge of the user's error rate—information typically unavailable in advance!

Experimental Evidence: Impact of Epsilon Mismatch

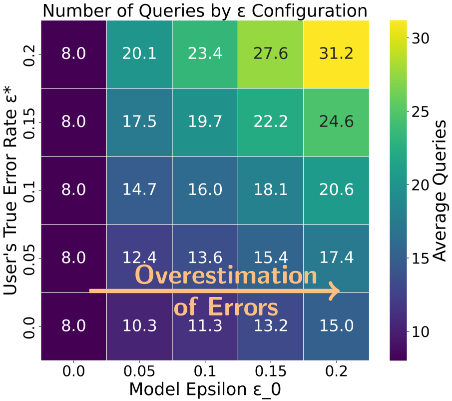


Underestimating errors degrades accuracy

Experimental Evidence: Impact of Epsilon Mismatch



Underestimating errors degrades accuracy



Overestimating errors increases query count

Learning the Error Rate: Joint Inference

Key Idea: Learn θ and ϵ simultaneously

Instead of fixing ϵ_0, treat it as an unknown parameter to be inferred

Fixed Error rate:

- ϵ_0 is fixed
- Update only $p(\theta|x, y)$
- Limited adaptability

Adaptive Error rate:

- ϵ is a random variable
- Update joint distribution $p(\theta, \epsilon|x, y)$
- Self-adjusts to actual error patterns

Likelihood:

$$p(y|x, \theta, \epsilon) = (1 - \epsilon) \cdot \delta(y, f(x, \theta)) + \epsilon \cdot q(y|x)$$

This is now a *parameterized family of likelihood functions* where each ϵ value defines a different likelihood model

BIG Algorithm with Joint Estimation

BIG with Fixed Error rate or Zero Error Model:

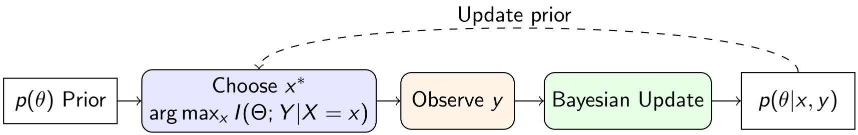


Figure 2: Bayesian update cycle: inference of θ with direct feedback

BIG with Adaptive Error rate Model:

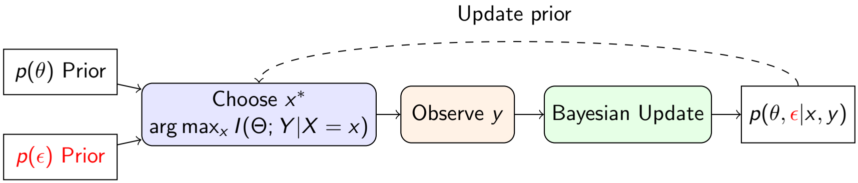
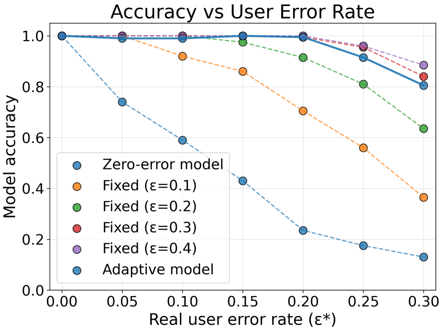
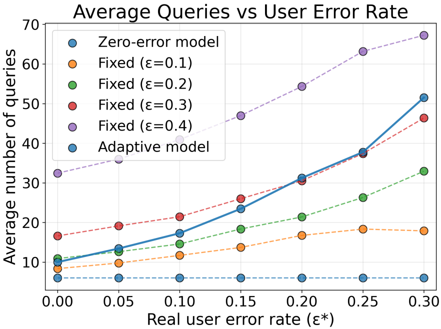


Figure 3: Bayesian update cycle: joint inference of θ and ϵ with direct feedback

Experimental Results: Adaptive Error rate Model Performance





Mathematical Continuity Between Error Models

Establishing a Formal Relationship Between Models

We will now demonstrate the mathematical continuity between our three error models:

1. **Zero Error** $\rightarrow$ **Fixed Error rate**
2. **Fixed Error rate** $\rightarrow$ **Adaptive Error rate**
3. **Complete Hierarchy**

Importance of Continuity: This continuity establishes that our three models form a coherent mathematical framework, where each model naturally extends from the previous one while preserving its essential properties.

Continuity Between Fixed and Zero Error rate Models

Proposition 1: Likelihood Continuity

The zero error model is a limiting case of the fixed error rate model as $\epsilon_0 \to 0$.

Continuity Between Fixed and Zero Error rate Models

Proposition 1: Likelihood Continuity

The zero error model is a limiting case of the fixed error rate model as $\epsilon_0 \to 0$.

Fixed Error rate Likelihood:

$$p(y|x, \theta, \epsilon_0) = (1 - \epsilon_0)\delta(y, f(x, \theta)) + \epsilon_0 \cdot q(y|x)$$

Zero Error Likelihood:

$$p(y|x, \theta) = \delta(y, f(x, \theta))$$

Likelihood Continuity:

$$\begin{aligned} \lim_{\epsilon_0 \to 0} p(y|x, \theta, \epsilon_0) &= \lim_{\epsilon_0 \to 0} [(1 - \epsilon_0)\delta(y, f(x, \theta)) + \epsilon_0 \cdot q(y|x)] \\ &= \delta(y, f(x, \theta)) = p(y|x, \theta) \end{aligned}$$

Continuity Between Adaptive and Fixed Error rate Models

Proposition 2: Likelihood Continuity

The fixed error rate model is a limiting case of the adaptive model as the distribution $p(\epsilon)$ approaches a Dirac delta at ϵ_0.

Continuity Between Adaptive and Fixed Error rate Models

Proposition 2: Likelihood Continuity

The fixed error rate model is a limiting case of the adaptive model as the distribution $p(\epsilon)$ approaches a Dirac delta at ϵ_0.

Adaptive model with discrete distribution on ϵ:

- Let $p_n(\epsilon)$ be a sequence of discrete distributions
- As $n \to \infty$, $p_n(\epsilon) \to \delta(\epsilon - \epsilon_0)$

Likelihood in adaptive model:

$$\begin{aligned} p(y|x, \theta, \mathcal{E}) &= \sum_{\epsilon} p(\epsilon) \cdot p(y|x, \theta, \epsilon) \\ &= \sum_{\epsilon} p(\epsilon) \cdot [(1-\epsilon)\delta(y, f(x, \theta)) + \epsilon \cdot q(y|x)] \end{aligned}$$

Limit as $p_n(\epsilon) \to \delta(\epsilon - \epsilon_0)$:

$$\lim_{n \to \infty} p(y|x, \theta, \mathcal{E}) = (1-\epsilon_0)\delta(y, f(x, \theta)) + \epsilon_0 \cdot q(y|x)$$

Continuity Between Models: Complete Picture

Model Hierarchy

Each model can be derived as a special case of the more general one:

- Zero error model: special case of fixed error rate model with $\epsilon_0 = 0$
- Fixed error rate model: special case of adaptive model with $p(\epsilon) = \delta(\epsilon - \epsilon_0)$

Significance

This continuity establishes a hierarchy of models where each generalizes the previous one:

$$\text{Zero Error} \subset \text{Fixed Error rate} \subset \text{Adaptive Error rate Model}$$

As special cases, the simpler models can be recovered from the more general ones.

Conclusion and Future Work

Summary of Contributions:

- Extended BIG framework to handle user errors
- Developed three models with increasing sophistication
- Proved mathematical continuity between models

Conclusion and Future Work

Summary of Contributions:

- Extended BIG framework to handle user errors
- Developed three models with increasing sophistication
- Proved mathematical continuity between models

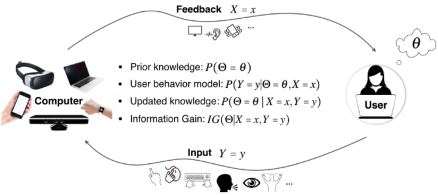
Future Directions:

- Using posterior distributions as priors for subsequent interactions, enabling continuous learning across multiple BIG instances
- Extending our discrete proofs to continuous distributions
- Exploring alternative utility functions beyond mutual information
- Validating the adaptive model in practical applications

Thank you for your attention!

Questions?

Contact:
hugo.miquel@telecom-paris.fr
julien.gori@sorbonne-universite.fr
olivier.rioul@telecom-paris.fr



Important Note on Continuous Distributions

Caution for Continuous ϵ Distributions

Our proof uses a discrete distribution for ϵ converging to a Dirac delta.

For continuous distributions:

- The integral form would be:

$$p(y|x,\theta,\mathcal{E}) = \int_0^1 p(\epsilon) \cdot p(y|x,\theta,\epsilon)\, d\epsilon$$

- Taking the limit requires exchanging limit and integration:

$$\lim_{n\to\infty} \int_0^1 p_n(\epsilon) \cdot p(y|x,\theta,\epsilon)\, d\epsilon = \int_0^1 \lim_{n\to\infty} p_n(\epsilon) \cdot p(y|x,\theta,\epsilon)\, d\epsilon$$

- This exchange requires additional assumptions and justification (uniform convergence, dominated convergence theorem, etc.)

References I

- Greenwade, George D. (1993). “The Comprehensive Tex Archive Network (CTAN)”. In: *TUGBoat* 14.3, pp. 342–351.
- Lee, Seung Won et al. (2023). “BIGaze: an eye-gaze action-guided Bayesian information gain framework for information exploration”. In: *Advanced Engineering Informatics* 58, p. 102159.
- Liu, Wanyu, Olivier Rioul, and Michel Beaudouin-Lafon (2021). “Bayesian Information Gain to Design Interaction”. In.
- — (2022). “Bayesian Information Gain to Design Interaction”. In: *Bayesian Methods for Interaction Design*.
- Liu, Wanyu et al. (2017). “Bignav: Bayesian information gain for guiding multiscale navigation”. In: *Proceedings of the 2017 CHI conference on human factors in computing systems*, pp. 5869–5880.
- Liu, Wanyu et al. (2018). “BIGFile: Bayesian information gain for fast file retrieval”. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, pp. 1–13.