Hierarchical Signal Detection Theory (HSDT)

Combining SDT and Hierarchical Models

Joachim Vandekerckhove

Winter 2025

Hierarchical SDT (HSDT)

HSDT: Why Combine Them?

Applying classical SDT to typical cognitive science data (multiple participants, conditions) faces challenges:

HSDT: Why Combine Them?

Applying classical SDT to typical cognitive science data (multiple participants, conditions) faces challenges:

1. Aggregated Analysis (Complete Pooling)

Sum counts across participants.

- × Ignores individual differences.
- × Can mislead (Simpsonś Paradox).
- × Violates SDT/Binomial assumptions.

HSDT: Why Combine Them?

Applying classical SDT to typical cognitive science data (multiple participants, conditions) faces challenges:

1. Aggregated Analysis (Complete Pooling)

Sum counts across participants.

- × Ignores individual differences.
- × Can mislead (Simpsonś Paradox).
- × Violates SDT/Binomial assumptions.

2. Individual Analysis (No Pooling)

Calculate d', c per person.

- × Unreliable with sparse data.
- × Extreme rates (0/1) $\implies$ infinite d'/c.
- × Needs ad-hoc corrections.
- × Inefficient (ignores population).
- × Hard to compare groups (noisy inputs).

HSDT: Why Combine Them?

Applying classical SDT to typical cognitive science data (multiple participants, conditions) faces challenges:

1. Aggregated Analysis (Complete Pooling)

Sum counts across participants.

- × Ignores individual differences.
- × Can mislead (Simpsonś Paradox).
- × Violates SDT/Binomial assumptions.

2. Individual Analysis (No Pooling)

Calculate d', c per person.

- × Unreliable with sparse data.
- × Extreme rates (0/1) $\implies$ infinite d'/c.
- × Needs ad-hoc corrections.
- × Inefficient (ignores population).
- × Hard to compare groups (noisy inputs).

HSDT Solution: Embed the SDT model within a hierarchical (often Bayesian) structure.

- Simultaneously estimate individual and population parameters.

HSDT: Bayesian Formulation (Example)

Assume P participants, K conditions.

1. **Likelihood (Level 1a - Data):** Binomial likelihood for Hits (H_{pk}) and False Alarms (FA_{pk}) for participant p in condition k, given trial counts ($N_{S,pk}, N_{N,pk}$) and latent probabilities ($\theta_{H,pk}, \theta_{FA,pk}$): $H_{pk} \sim \text{Binomial}(N_{S,pk}, \theta_{H,pk})$ $FA_{pk} \sim \text{Binomial}(N_{N,pk}, \theta_{FA,pk})$

HSDT: Bayesian Formulation (Example)

Assume P participants, K conditions.

1. **Likelihood (Level 1a - Data):** Binomial likelihood for Hits (H_{pk}) and False Alarms (FA_{pk}) for participant p in condition k, given trial counts ($N_{S,pk}, N_{N,pk}$) and latent probabilities ($\theta_{H,pk}, \theta_{FA,pk}$): $H_{pk} \sim \text{Binomial}(N_{S,pk}, \theta_{H,pk})$ $FA_{pk} \sim \text{Binomial}(N_{N,pk}, \theta_{FA,pk})$
2. **SDT Link (Level 1b - Transformation):** Link probabilities to participant-condition SDT parameters (d'_{pk}, c_{pk}) via standard normal CDF *ormalCDF* (assuming equal-variance Gaussian SDT): $\theta_{H,pk} = ormalCDF d'_{pk}/2 - c_{pk}$ $\theta_{FA,pk} = ormalCDF - d'_{pk}/2 - c_{pk}$

HSDT: Bayesian Formulation (Example)

Assume P participants, K conditions.

1. **Likelihood (Level 1a - Data):** Binomial likelihood for Hits (H_{pk}) and False Alarms (FA_{pk}) for participant p in condition k, given trial counts ($N_{S,pk}, N_{N,pk}$) and latent probabilities ($\theta_{H,pk}, \theta_{FA,pk}$): $H_{pk} \sim \text{Binomial}(N_{S,pk}, \theta_{H,pk})$
   $FA_{pk} \sim \text{Binomial}(N_{N,pk}, \theta_{FA,pk})$
2. **SDT Link (Level 1b - Transformation):** Link probabilities to participant-condition SDT parameters (d'_{pk}, c_{pk}) via standard normal CDF $ormalCDF$ (assuming equal-variance Gaussian SDT): $\theta_{H,pk} = ormalCDF d'_{pk}/2 - c_{pk}$
   $\theta_{FA,pk} = ormalCDF{-d'_{pk}}/2 - c_{pk}$
3. **Individual Parameters (Level 2 - Random Effects):** Individual parameters (d'_{pk}, c_{pk}) drawn from condition-specific population distributions: $d'_{pk} \sim \mathcal{N}(\mu_{d'_k}, \sigma^2_{d'_k})$
   $c_{pk} \sim \mathcal{N}(\mu_{c_k}, \sigma^2_{c_k})$
   (Optionally, use Multivariate Normal for correlation $\rho_{dc,k}$)

HSDT: Bayesian Formulation (Example)

Assume P participants, K conditions.

1. **Likelihood (Level 1a - Data):** Binomial likelihood for Hits (H_{pk}) and False Alarms (FA_{pk}) for participant p in condition k, given trial counts ($N_{S,pk}, N_{N,pk}$) and latent probabilities ($\theta_{H,pk}, \theta_{FA,pk}$): $H_{pk} \sim \text{Binomial}(N_{S,pk}, \theta_{H,pk})$
   $FA_{pk} \sim \text{Binomial}(N_{N,pk}, \theta_{FA,pk})$
2. **SDT Link (Level 1b - Transformation):** Link probabilities to participant-condition SDT parameters (d'_{pk}, c_{pk}) via standard normal CDF $ormalCDF$ (assuming equal-variance Gaussian SDT): $\theta_{H,pk} = ormalCDF d'_{pk}/2 - c_{pk}$
   $\theta_{FA,pk} = ormalCDF - d'_{pk}/2 - c_{pk}$
3. **Individual Parameters (Level 2 - Random Effects):** Individual parameters (d'_{pk}, c_{pk}) drawn from condition-specific population distributions: $d'_{pk} \sim \mathcal{N}(\mu_{d'_k}, \sigma^2_{d'_k})$
   $c_{pk} \sim \mathcal{N}(\mu_{c_k}, \sigma^2_{c_k})$
   (Optionally, use Multivariate Normal for correlation $\rho_{dc,k}$)
4. **Population Parameters (Level 3 - Fixed Effects / Predictors):** Model the condition means ($\mu_{d'_k}, \mu_{c_k}$) and standard deviations ($\sigma_{d'_k}, \sigma_{c_k}$). Can be estimated per condition or modeled (e.g., via regression): $\mu_{d'_k} = \beta_{0d} + \beta_{1d} \times \text{Predictor}_k + \ldots$
   $\mu_{c_k} = \beta_{0c} + \beta_{1c} \times \text{Predictor}_k + \ldots$

HSDT: Bayesian Formulation (Example)

Assume P participants, K conditions.

1. **Likelihood (Level 1a - Data):** Binomial likelihood for Hits (H_{pk}) and False Alarms (FA_{pk}) for participant p in condition k, given trial counts ($N_{S,pk}, N_{N,pk}$) and latent probabilities ($\theta_{H,pk}, \theta_{FA,pk}$): $H_{pk} \sim \text{Binomial}(N_{S,pk}, \theta_{H,pk})$
   $FA_{pk} \sim \text{Binomial}(N_{N,pk}, \theta_{FA,pk})$
2. **SDT Link (Level 1b - Transformation):** Link probabilities to participant-condition SDT parameters (d'_{pk}, c_{pk}) via standard normal CDF $ormalCDF$ (assuming equal-variance Gaussian SDT): $\theta_{H,pk} = ormalCDF d'_{pk}/2 - c_{pk}$
   $\theta_{FA,pk} = ormalCDF -d'_{pk}/2 - c_{pk}$
3. **Individual Parameters (Level 2 - Random Effects):** Individual parameters (d'_{pk}, c_{pk}) drawn from condition-specific population distributions: $d'_{pk} \sim \mathcal{N}(\mu_{d'_k}, \sigma^2_{d'_k})$
   $c_{pk} \sim \mathcal{N}(\mu_{c_k}, \sigma^2_{c_k})$
   (Optionally, use Multivariate Normal for correlation $\rho_{dc,k}$)
4. **Population Parameters (Level 3 - Fixed Effects / Predictors):** Model the condition means ($\mu_{d'_k}, \mu_{c_k}$) and standard deviations ($\sigma_{d'_k}, \sigma_{c_k}$). Can be estimated per condition or modeled (e.g., via regression): $\mu_{d'_k} = \beta_{0d} + \beta_{1d} \times \text{Predictor}_k + \ldots$
   $\mu_{c_k} = \beta_{0c} + \beta_{1c} \times \text{Predictor}_k + \ldots$
5. **Hyperpriors:** Priors on all top-level parameters (e.g., βś, or μś, σś if Level 3 is simpler)

HSDT: Advantages

- **Handles Extreme Rates Gracefully:** Bayesian estimation with priors naturally handles HR/FAR of 0 or 1 without ad-hoc corrections. Posterior estimates remain finite and interpretable.

HSDT: Advantages

- **Handles Extreme Rates Gracefully:** Bayesian estimation with priors naturally handles HR/FAR of 0 or 1 without ad-hoc corrections. Posterior estimates remain finite and interpretable.
- **Improved Precision:** Partial pooling leads to more precise and stable estimates of both individual and group-level parameters.

HSDT: Advantages

- **Handles Extreme Rates Gracefully:** Bayesian estimation with priors naturally handles HR/FAR of 0 or 1 without ad-hoc corrections. Posterior estimates remain finite and interpretable.
- **Improved Precision:** Partial pooling leads to more precise and stable estimates of both individual and group-level parameters.
- **Quantifies Individual Differences:** Directly estimates the magnitude of between-participant variability ($\sigma_{d'}$, σ_c) and how it might differ across conditions.

HSDT: Advantages

- **Handles Extreme Rates Gracefully:** Bayesian estimation with priors naturally handles HR/FAR of 0 or 1 without ad-hoc corrections. Posterior estimates remain finite and interpretable.
- **Improved Precision:** Partial pooling leads to more precise and stable estimates of both individual and group-level parameters.
- **Quantifies Individual Differences:** Directly estimates the magnitude of between-participant variability ($\sigma_{d'}$, σ_c) and how it might differ across conditions.
- **Unified Inference:** Simultaneous inference on individual effects, group averages, condition differences, and variability within one coherent model.

HSDT: Advantages

- **Handles Extreme Rates Gracefully:** Bayesian estimation with priors naturally handles HR/FAR of 0 or 1 without ad-hoc corrections. Posterior estimates remain finite and interpretable.
- **Improved Precision:** Partial pooling leads to more precise and stable estimates of both individual and group-level parameters.
- **Quantifies Individual Differences:** Directly estimates the magnitude of between-participant variability ($\sigma_{d'}$, σ_c) and how it might differ across conditions.
- **Unified Inference:** Simultaneous inference on individual effects, group averages, condition differences, and variability within one coherent model.
- **Model Flexibility & Comparison:** Easily incorporates covariates (e.g., age). Principled comparison of different SDT assumptions (e.g., equal vs. unequal variance) or hierarchical structures using Bayesian tools (PPCs, LOO-CV).

HSDT: Implementation Tips

- **Parameterization:** Be clear about the definition of c (relative vs. absolute). While (d', c) is standard, others exist.

HSDT: Implementation Tips

- **Parameterization:** Be clear about the definition of c (relative vs. absolute). While (d', c) is standard, others exist.
- **Priors:** Choice of priors, especially for variance components ($\sigma\acute{s}$), matters. Affects shrinkage. Use weakly informative priors (e.g., Half-Normal, Half-Cauchy) informed by domain knowledge and perform sensitivity analysis.

HSDT: Implementation Tips

- **Parameterization:** Be clear about the definition of c (relative vs. absolute). While (d', c) is standard, others exist.
- **Priors:** Choice of priors, especially for variance components (σś), matters. Affects shrinkage. Use weakly informative priors (e.g., Half-Normal, Half-Cauchy) informed by domain knowledge and perform sensitivity analysis.
- **Sampling Efficiency (Non-Centered Parameterization): Often Essential!** Improves MCMC performance drastically, especially needed if group variances (σś) might be small.
  - Instead of: $d'_{pk} \sim \mathcal{N}(\mu_{d'_k}, \sigma^2_{d'_k})$
  - Use: $z_{d',pk} \sim \mathcal{N}(0, 1)$ and define $d'_{pk} = \mu_{d'_k} + z_{d',pk} \times \sigma_{d'_k}$.
  - (Similarly for c_{pk})
  - Decouples individual effects from group parameters in sampling.

HSDT: Implementation Tips

- **Parameterization:** Be clear about the definition of c (relative vs. absolute). While (d', c) is standard, others exist.
- **Priors:** Choice of priors, especially for variance components (σś), matters. Affects shrinkage. Use weakly informative priors (e.g., Half-Normal, Half-Cauchy) informed by domain knowledge and perform sensitivity analysis.
- **Sampling Efficiency (Non-Centered Parameterization): Often Essential!** Improves MCMC performance drastically, especially needed if group variances (σś) might be small.
    - Instead of: $d'_{pk} \sim \mathcal{N}(\mu_{d'_k}, \sigma^2_{d'_k})$
    - Use: $z_{d',pk} \sim \mathcal{N}(0,1)$ and define $d'_{pk} = \mu_{d'_k} + z_{d',pk} \times \sigma_{d'_k}$.
    - (Similarly for c_{pk})
    - Decouples individual effects from group parameters in sampling.
- **Model Checking: Critical!**
    - MCMC diagnostics ($\hat{R}$, ESS, divergences, trace plots).
    - Posterior Predictive Checks (PPCs): Do simulated datasets look like the real data

HSDT: Interpretation

Focus on posterior distributions of key parameters:

- **Population Averages (e.g., $\mu_{d'_k}, \mu_{c_k}$ or βś):** Estimated average sensitivity/bias in each condition. Credible differences between conditions (e.g., posterior of $\mu_{d'_1} - \mu_{d'_2}$).

HSDT: Interpretation

Focus on posterior distributions of key parameters:

- **Population Averages (e.g., $\mu_{d'_k}, \mu_{c_k}$ or βś):** Estimated average sensitivity/bias in each condition. Credible differences between conditions (e.g., posterior of $\mu_{d'_1} - \mu_{d'_2}$).
- **Population Variability (e.g., $\sigma_{d'_k}, \sigma_{c_k}$):** How much do individuals typically differ? Is variability different across conditions?

HSDT: Interpretation

Focus on posterior distributions of key parameters:

- **Population Averages (e.g., $\mu_{d'_k}, \mu_{c_k}$ or βś):** Estimated average sensitivity/bias in each condition. Credible differences between conditions (e.g., posterior of $\mu_{d'_1} - \mu_{d'_2}$).
- **Population Variability (e.g., $\sigma_{d'_k}, \sigma_{c_k}$):** How much do individuals typically differ? Is variability different across conditions?
- **Individual Estimates (e.g., posterior mean/median of d'_{pk}, c_{pk}):** Best estimates for specific individuals/conditions (reflecting shrinkage). More reliable than non-hierarchical estimates, especially with sparse data.

HSDT: Interpretation

Focus on posterior distributions of key parameters:

- **Population Averages (e.g., $\mu_{d'_k}, \mu_{c_k}$ or βś):** Estimated average sensitivity/bias in each condition. Credible differences between conditions (e.g., posterior of $\mu_{d'_1} - \mu_{d'_2}$).
- **Population Variability (e.g., $\sigma_{d'_k}, \sigma_{c_k}$):** How much do individuals typically differ? Is variability different across conditions?
- **Individual Estimates (e.g., posterior mean/median of d'_{pk}, c_{pk}):** Best estimates for specific individuals/conditions (reflecting shrinkage). More reliable than non-hierarchical estimates, especially with sparse data.
- **Correlations (e.g., $\rho_{dc,k}$):** Systematic relationship between sensitivity and bias across participants?

HSDT: Interpretation

Focus on posterior distributions of key parameters:

- **Population Averages (e.g., $\mu_{d'_k}, \mu_{c_k}$ or βś):** Estimated average sensitivity/bias in each condition. Credible differences between conditions (e.g., posterior of $\mu_{d'_1} - \mu_{d'_2}$).
- **Population Variability (e.g., $\sigma_{d'_k}, \sigma_{c_k}$):** How much do individuals typically differ? Is variability different across conditions?
- **Individual Estimates (e.g., posterior mean/median of d'_{pk}, c_{pk}):** Best estimates for specific individuals/conditions (reflecting shrinkage). More reliable than non-hierarchical estimates, especially with sparse data.
- **Correlations (e.g., $\rho_{dc,k}$):** Systematic relationship between sensitivity and bias across participants?
- **Covariate Effects:** How do covariates (if included) relate to performance or individual differences?

HSDT: Example Analysis Output

Hierarchical Bayesian models allow rich visualization of results. (Example plots below would be generated from running models like ‘src/sdt/hierarchical.py‘ or ‘src/sdt/full.py‘)

HSDT: Example Analysis Output

Hierarchical Bayesian models allow rich visualization of results. (Example plots below would be generated from running models like ‘src/sdt/hierarchical.py‘ or ‘src/sdt/full.py‘)

Population Parameters:

- Posterior distributions for group means (e.g., $\mu_{d'}$, μ_c).
- Posterior distributions for group SDs (e.g., $\sigma_{d'}$, σ_c).
- Posterior for difference between conditions.

Placeholder:
Posterior plot for $\mu_{d'}$

Placeholder:

HSDT: Example Analysis Output

Hierarchical Bayesian models allow rich visualization of results. (Example plots below would be generated from running models like ‘src/sdt/hierarchical.py‘ or ‘src/sdt/full.py‘)

Population Parameters:

- Posterior distributions for group means (e.g., $\mu_{d'}$, μ_c).
- Posterior distributions for group SDs (e.g., $\sigma_{d'}$, σ_c).
- Posterior for difference between conditions.

Placeholder:
Posterior plot for $\mu_{d'}$

Placeholder:

Individual Parameters:

- Forest plots showing individual participant estimates (e.g., d'_p, c_p) with uncertainty intervals.
- Demonstrates shrinkage towards the group mean.

Placeholder:
Forest plot for individual d'_p

HSDT: Key References i

- Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press. *(Provides clear explanations and code examples for HSDT, highly recommended).*
- Rouder, J. N., & Lu, J. (2005). An introduction to Bayesian hierarchical models with an application in the theory of signal detection. *Psychonomic Bulletin & Review*, *12*(4), 573–604. *(The foundational paper clearly motivating and outlining the HSDT approach).*
- Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan* (2nd ed.). Academic Press. *(Provides the general Bayesian hierarchical framework and techniques applicable here).*
- DeCarlo, L. T. (2010). The analysis of signal detection data using a hierarchical Bayesian random-coefficient model. *British Journal of Mathematical and Statistical Psychology*, *63*(Pt 1), 69–95. *(Focuses on random-coefficient modeling for SDT).*

HSDT: Key References ii

- Morey, R. D., Rouder, J. N., Verhagen, J., & Wagenmakers, E. J. (2008). Computation tool for fitting hierarchical signal detection models. *Behavior Research Methods*, *40*(1), 9-14. *(Discusses software implementation).*
- Wright, D. B., Horry, R., & Skagerberg, E. M. (2009). Functions for traditional and multilevel approaches to signal detection theory. *Behavior Research Methods*, *41*(1), 1-10. *(Discusses practical calculation and multilevel aspects).*
- See also references from Hierarchical Modeling lecture.

Conclusion

The integration of SDT and hierarchical modeling represents a mature and powerful approach for analyzing decision-making data with individual differences and complex structures.

Conclusion

The integration of SDT and hierarchical modeling represents a mature and powerful approach for analyzing decision-making data with individual differences and complex structures. **Key Takeaways:**

1. **Theoretical Depth:** SDT decomposes performance into sensitivity and bias.
2. **Statistical Rigor:** Hierarchical models handle non-independence via partial pooling.
3. **HSDT Synergy:** Principled study of individual/group differences in SDT parameters.
4. **Bayesian Advantages:** Coherent uncertainty, computation (MCMC), model checking/comparison.
5. **Practical Relevance:** Essential for analyzing typical psychology data, providing reliable and interpretable results.

Conclusion

The integration of SDT and hierarchical modeling represents a mature and powerful approach for analyzing decision-making data with individual differences and complex structures. **Key Takeaways:**

1. **Theoretical Depth:** SDT decomposes performance into sensitivity and bias.
2. **Statistical Rigor:** Hierarchical models handle non-independence via partial pooling.
3. **HSDT Synergy:** Principled study of individual/group differences in SDT parameters.
4. **Bayesian Advantages:** Coherent uncertainty, computation (MCMC), model checking/comparison.
5. **Practical Relevance:** Essential for analyzing typical psychology data, providing reliable and interpretable results.

Mastery of HSDT allows asking more sophisticated questions and drawing robust

Hierarchical Signal Detection Theory (HSDT)

Combining SDT and Hierarchical Models

Joachim Vandekerckhove

Winter 2025