Bayesian Cognitive Modeling applied to Signal Detection Theory: the Mirror Effect in a Perceptual Task as a special guest

Adriana F. Chávez-De la Peña, Michael D. Lee

September, 2019

Abstract:
Signal Detection Theory (SDT) has provided one of the most well-known and broadly applied statistical models within Cognitive Sciences to study a wide range of phenomena where decision-making systems depend on the detection of specific stimuli presented within their noisy environment to produce an optimal and adaptive response. Within Recognition Memory studies where SDT is applied to compare subjects' performance between two classes of stimuli that are known to be deferentially recognized, it has consistently been found that this difference is reflected in the identification of both target and lure stimuli, as measured by hit and false alarm rates in Signal Detection Theory. The implied order of the signal and noise distributions has led to the identification of this pattern of response as "the Mirror Effect". Since this phenomena has been predominantly studied within recognition memory tasks, most attempts to account for it theorize about high-level processes engaged in the study phase. To test the generalizability of the Mirror Effect to other domains where signal detection theory has been applied, we designed a perceptual task with two levels of discriminability defined using an optical illusion. By conducting a step by step replication of the mean-based analysis reported in the literature, we present evidence of the mirror effect outside recognition memory, and then, we continue to present a more detailed model based analysis, using Bayesian methods to assess the existence of the mirror effect at both the group and individual level.

Keywords: Signal Detection Theory; Bayesian Modeling; Perception; Recognition Memory; Mirror Effect; Ebbinghaus Illusion.

1 Introduction

We live in a world full of changes, noisy information and uncertainty. Every decision-making system is constantly exposed to a huge variety of stimuli that does not always provide relevant information for their adaptation. Once that organisms learn the most essential contingency rules operating in their environment, the detection of specific stimuli (i.e. signals) among all non-informative stimuli (i.e. noise), becomes a priority task in order for them to be able to adapt to the current state of the world. We can think of a huge variety of examples to illustrate the importance of this signal-detection kind of tasks, from a vulnerable animal that has to determine whether or not the sound it has just heard in the woods represents a threat or not to decide between running and keep gathering food, to the physician that has to determine if the tomography scan she's inspecting contains evidence of a cancerous tumor or not, in order to produce a diagnostic that could have a strong effect on the patient's lifestyle.

Signal Detection Theory (SDT) has provided a very intuitive framework to understand this kind of situations and that has been widely applied within and outside Psychological Science to account for a huge variety of phenomena involving this kind of dilemmas.

1.1 Signal Detection Theory Model

Signal Detection Theory appeared for the first time in 1954 and just like many other scientific and technological developments born around that time, its purpose was to contribute to the solution of a necessity derived from the Second World War: the development and improvement of radars (Peterson, Birdsall y Fox, **?**). Not long after that, this theoretical framework was brought into the Psychological domain to study organic perception (Tanner y Swets, **?**; Swets, Tanner y Birdsall, **?**).

The SDT model works as a descriptive model that captures the problem faced by any decision-making system who finds itself in the necessity to decide whether or not a signal is present within its immediate environment, in order to produce the most appropriate response to the contingencies its presence would be announcing, (Tanner y Swets, **?**; Swets, **?**).

The core contribution of the model is that it embodies the idea that there's always a certain level of uncertainty involved in this kind of tasks. This notion is represented by using probability distributions (typically, normal distributions) to describe the variability that both signal and noise stimuli present in terms of the values with which they can be associated along a "decision axis" that contains whatever kind of information the system is sensing and processing to produce a detection judgement ("Yes, the signal is there" / "No, the signal is absent"). This variability can be explained either by assuming stimuli vary in the way they appear in the world (e.g. there's a lot of variability in terms of animals that could be considered a threat) or because there's always variability in the way decision-making systems read their environment (as observed in Psychophysics). Uncertainty comes from the notion that a portion of the noise distribution overlaps the signal distribution, allowing for the existence of a range of values on the decision axis that can be associated with either kind of stimuli. This is what makes the task complicated: decision-making systems cannot rely blindly on the evidence they perceive to emit a judgement, (Killeen, **?**; Wickens **?**).

When organisms are presented with a stimulus located at the overlapping area of the signal and noise distributions, they have to decide which binary judgment to make, by weighting the information they're currently receiving from the environment with all the information available about its structure, which can be classified in two big categories: 1) Information about the probabilistic nature of the environment and 2) Information regarding the consequences at risk, (Killeen, **?**). Let's think again about the physician who's evaluating some tomography scans: if she's not confident enough about the information being presented to diagnose the presence of a certain disease, she would have to take into account the information she has about the probability at which this disease is usually presented and the risks involved with letting a sick patient go without a treatment versus treating someone who's not sick with a very invasive procedure.

The SDT model is a decision model that assumes that binary detection judgments are produced by comparing the evidence that's currently being observed with a decision criterion which cuts the "decision axis" and which location is determined by the decision-making system. Thus, the probability of making a "Hit" or a "False alarm" depend on which portion of the signal and noise distributions fall above the criterion and the probability of making a "Correct rejection" or a "Miss" is related to how much is left below it, (See Figure **??**)

One of the most valued features of the SDT model is the way in which it allows to measure separately the influence of two great factors on the decision-making system's performance: The discriminability of the signal stimuli from noise, (as represented by the distance between the mean of the signal and the noise distribution, captured by the d' parameter), and the system's bias towards a specific binary detection judgment over the other. SDT provides two distinct measures of bias: the likelihood ratio between the points at which the criterion cuts the signal and the noise distribution (the β parameter) and the distance between the observed criterion and the ideal location of the criterion assuming no bias at the point where the two distributions overlap (measured by a parameter typically identified as c).

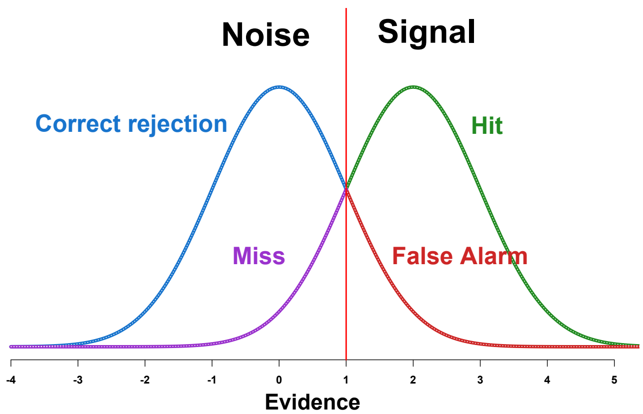


Figure 1: Graphical representation of the Signal Detection Theory model

To this day, SDT is one of the best-known models in Psychological Science. It has been applied to the study of a huge variety of phenomena, both as a tool for data analysis and as a cognitive model to describe the cognitive processes behind the emission of binary choices being made under uncertainty. One can find examples of its application to the study of perception (Rosenholtz, **?**; Pessoa, Japee y Ungerleider, **?**; Wallis y Horswill, **?**), where it was first adopted within Psychology; the emission of clinical diagnosis (Grossberg y Grant, **?**; Swets, Dawes y Monahan, **?**; Boutis, Pecaric, Seeto y Pusic, **?**); the syntomathology associated with all kinds of clinical conditions (Westermann y Lincoln, **?**; Bonnel y cols., **?**; Brown, Kosslyn, Breiter, Baer y Jenike, **?**; Naliboff y Cohen, **?**); eyewitness testimonies (Gronlund, Wixted y Mickes, **?**; Wixted y Mickes, **?**; Wixted, Miches, Dunn, Clark y Wells, **?**), and an incredibly ongroing etcetera (Gordon y Clark, **?**; Nuechterlein, **?**; Harvey Jr., Hammond, Lusk y Mross, **?**; Verghese, **?**).

The huge number of publications on the SDT is not limited to its application to different phenomena, its utility has motivated the production of several tutorials and manuals, which are aimed to provide either an introduction to the benefits of the model (Killeen, **?**); a detailed description of its assumptions and implications (McNicol, **?**; Wickens **?**); or a guide to conduct data analysis under the SDT framework, (Stainslaw y Todorov, **?**).

1.2 The mirror effect

Among the previously described variety of phenomena where SDT has been applied, its relevance to the study of Recognition Memory is definitely one of the most notorious. In a typical recognition memory task, participants are presented with a set of stimuli that they are either explicitly asked to study or with which they are expected to interact in any given way; this first phase is typically referred as the Study phase and it can involve an intentional or an incidental study of the stimuli presented. After some time has passed, participants are faced again with the same stimuli they experienced on the first phase plus another set of stimuli that hasn't been used before. The recognition memory task, as its name clearly suggest, requires subjects to identify the stimuli that they had encountered in the first phase (the signal) from those that are totally new (the noise).

When SDT has been used to analyze data obtain from Recognition Memory tasks where two different classes of stimuli (where one is known to be more easily recognized (the A class) than the other (the B class)) are presented without subjects being aware of it, and their performance is analyze separately between the A and the B class, a consistent pattern of response has been reported, showing that not only more Hits are being made on the A class, but also less False alarms. This particular phenomenon has been identified in the literature as the Mirror Effect, given the implications it has in terms of the underlying signal and noise distributions, (Glanzer y Bowles, **?**; Glanzer, Adams, Iverson y Kim, **?**). Evidence for the Mirror Effect has been reported across different procedures typically used to collect detection judgements (the "Yes/No" task; two alternative forced choice and confidence ratings) and a wide variety of variables used to define the A and B classes. (Glanzer y Adams, **?**).

The Mirror Effect has only been studied within Recognition Memory literature and thus, the majority of the models and theories proposed to account for it tend to do it in terms of high-level processes involved in the study phase, (Glanzer, Adams, Iverson y Kim, **?**; Glanzer, Kim y Adams, **?**; Glanzer, Hilford y Maloney, **?**), as if it was a reflection of the inner cognitive processes engaged in recognition memory tasks.

The data reported in the present study provides evidence for the existence of the Mirror Effect outside Recognition Memory, within a merely perceptual task where two classes of stimuli were proposed based on the literature on Optical Illusions (in particular, the Ebbinghaus illusion). This study aims to explore the generalizability of the Mirror Effect while providing a more detailed analysis of this phenomenon at the individual level and its implications in terms of the SDT model, by means of the application of Bayesian methods.

2 Experiments

To evaluate the generalizability of the Mirror Effect, we proposed a binary perceptual detection task where participants had to indicate whether two circles presented for comparison on screen had the same size (signal trials) or not (noise trials). The same task was used in two different settings, referred as Experiment 1 and Experiment 2. On Experiment 1, one of these two circles was used as the center circle of an Ebbinghaus illusion, while in Experiment 2 both circles were the central circles of their own Ebbinghaus illusion. Details about the stimuli design are discussed in a later section.

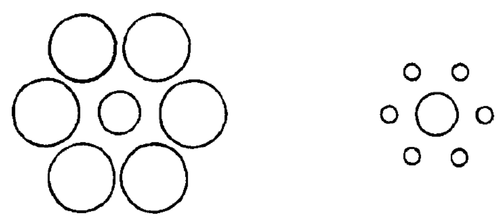


Figure 2: Illustration of the Ebbinghaus Illusion. The center circles have the same size, (illustration taken from Massaro y Anderson, **?**)

The Ebbinghaus Illusion (see Fig **??**), is a well-known and widely studied optical illusion that has been typically explained as the result of a perceptual mechanism that processes and estimates sizes based on its contrast witb regards of the surrounding stimuli. This illusion tends to be illustrated with the Ebbinghaus figures (also known as the Titchener's circles), where the size of an inner circle is either overestimated or underestimated due to its relation to a set of surrounding circles which are uniformally bigger (Underestimation effect), or smaller (Overestimation effect), (Coren, **?**; Coren y Miller, **?**; De Fockert, Davidoff, Fagot, Parron y Goldstein, **?**).

We used what has been reported in the literature about the variables that have an impact on the intensity of the Ebbinghaus illusion (Massaro y Anderson, **?**; Girgus, Coren y Agdern, **?**; Roberts, HArris y Yates, **?**) to construct two levels of discriminability that would allow us to make our task analogous to those recognition tasks where the Mirror Effect has been reported. Figure **??** summarizes the results obtained in a study where the number (x-axis) and size (y-axis) of the external circles of several Ebbinghaus figures were manipulated to assess their effect on the intensity of the illusion elicited. This graph presents the mean estimations registered for the diameter of the central circle of the figures presented, and so, it can be observed that the discrepancy between the real diameter and participants' estimation increased as a function of the number of external circles included in the figures and, according to the vertical distance between every line, that the intensity of the illusion is proportional to the difference between the size of the external circles and the center circle, for either the Overestimation or the Underestimation effect, (Massaro y Anderson, ?).

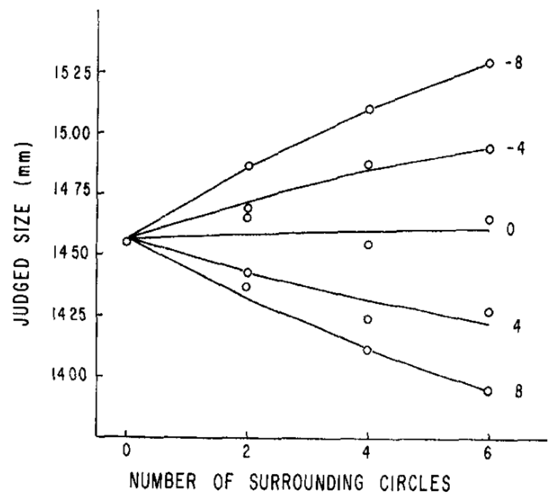


Figure 3: Results obtained in an experiment where the number and size of the external circles presented on different Ebbinghaus figures were manipulated to assess their effect on the illusion evoked, (Massaro y Anderson, ?).

According to Massaro and Anderson's findings (?), for our perceptual task we decided to construct the A class and the B class stimuli as follows:

- **Class A ("Higher discriminability"):** Ebbinghaus figures with few external circles (with two levels: 2 or 3 external circles)
- **Class B ("Lower discriminability"):** Ebbinghaus figures with more external circles (with two levels: 7 or 8 external circles)

For all figures constructed, external circles meant to induce an Overestimation effect were presented with a diameter of $0.5cm$ while external circles meant to induce an Underestimation effect had a constant diameter of $6cm$, these diameters were proposed so that they were at least $2x$ bigger and $0.5x$ smaller than the size of the central circles (which could vary from $1cm$ to $3cm$, with increments of $0.5cm$). The distance between the external and the central circles was not fixed nor controlled across the different figures created.

2.1 Experiment 1

On Experiment 1, participants had to compare the size of a constant stimulus (a circle with a diameter of $2cm$, which location on the left side of the screen remained fixed across trials) with the size of the central circle of an Ebbinghaus figure displayed on the right side of the screen.

Stimuli presented on Experiment 1 were constructed following a 5x2x2 factorial design, illustrated in Figure **??**, for each class of stimuli A and B. The reference circle with a constant diameter of $2cm$ was compared with the center circle of an Ebbinghaus illusion which could present five diameter lengths (from $1cm$ to $3cm$ in increments of $0.5cm$), with the external circles inducing either an Overestimation (external circles with diameter of $0.5cm$) or Underestimation (external circles with a diameter of $6cm$) effect and two different levels of number of external circles per class (2 or 3 external circles for the A class and 7 or 8, for the B class). We designed 16 different "Noise stimuli" and 4 different "Signal stimuli" for each class of stimuli. To account for the variability of choices made by participants, and to make both signal and noise trials equal in terms of the number of times they were presented, we repeated each of the 16 noise stimuli 10 times, while presenting each of the 4 signal stimuli 40 times. This lead to a total of 160 noise trials and 160 signal trials per class of stimuli, with a total of 640 trials for the entire experiment.

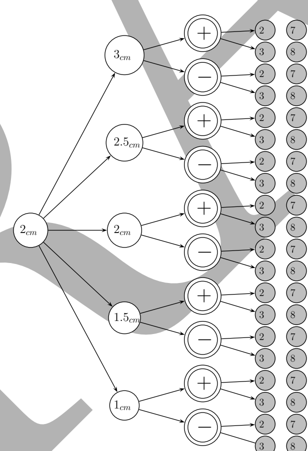


Figure 4: Stimuli design for Experiment 1: A 5x2x2 factorial design was used to construct the Ebbinghaus figures to be presented with five different diameters for the central circle, two different diameters for the external circles to produce an Overestimation or Underestimation effect and two levels of "Number of external circles" per class.

2.2 Experiment 2

On Experiment 2, the main detection task remained the same "Are the two central circles the same size?". However, it differ from Experiment 1 because this time, the two circles that participants were asked to compare in sized were used as the central circle of their own Ebbinghaus figure.

A diferencia del Experimento 1, donde uno de los círculos a comparar era constante, en el Experimento 2 se varió el diámetro de los dos círculos a comparar. Para ello se ello se utilizaron los mismos cinco tamaños de círculo central (de 1 a 3 cm en saltos de 0.5 cm) y por lo tanto, con cinco combinaciones posibles para las Parejas-señal. Así mismo, se formaron cinco Parejas-ruido juntando arbitrariamente valores de círculo central que guardasen una diferencia de 0.5 cm entre sí -1 vs 1.5; 1.5 vs 2; 2 vs 2.5 y 2.5 vs 3 cm- con una quinta pareja con una diferencia de 1 cm entre los valores de círculo central intermedios -1.5 cm vs 2.5 cm-. Por cada una de estas 10 parejas, se crearon cuatro variaciones por condición, de acuerdo con las combinaciones posibles de niveles de *'número de círculos externos'* (2 círculos externos a ambos lados, 3 círculos externos a ambos lados, 2 en izquierdo y 3 en derecho, y 3 en izquierdo y 2 en derecho en la condición fácil; 7 círculos externos a ambos lados, 8 círculos externos ambos lados, 7 círculos del lado izquierdo y 8 en el derecho y 8 círculos en el lado izquierdo y 7 en el derecho en la condición difícil). En total, el Experimento 2 estuvo compuesto por 80 parejas diferentes de figuras de Ebbinghaus cuyos círculos centrales debían compararse, 40 con la señal y 40 con el ruido y 20 de cada uno por condición.

Each of the 80 pairs proposed for comparison, Cada una de las 80 parejas diseñadas para el Experimento 2 se presentó 8 veces, en cuatro colores diferentes (púrpura, anaranjado, azul y verde) para prevenir la fatiga de los participantes, contrabalanceando la posición de las ilusiones de sobrestimación y subestimación a la derecha o izquierda de la pantalla. De tal forma que el Experimento 2 estuvo compuesto por un total de 640 ensayos, 320 por cada tipo de ensayo (ruido y señal) y 160 por cada clase.

2.3 Materials

The experiment was programmed and executed in **PsychoPy v.12** (Peirce, **?**), using a Mac computer ($59.5cm$ x $34cm$ screen). Experimental sessions were ran inside an isolated room, with participants being sat in a fixed chair $1.10m$ away from the screen.

2.4 Participants

A total of forty-one students (ages 18 to 21) of the Psychology School of the National Autonomous University of Mexico participated in one of the two ex-

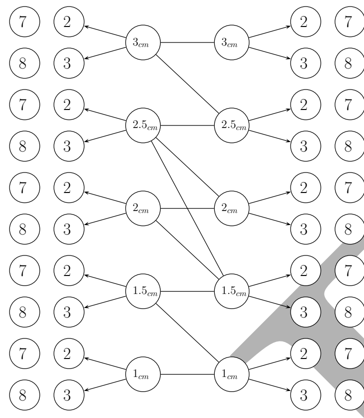


Figure 5: Stimuli design for Experiment 2:

periments conducted (20 in Experiment 1 and 41 in Experiment 2). Both experiments were conducted during the same period of time, with participants being assigned to one of these, without them knowing that there was more than one Experiment.

2.5 Procedure

The detection task was conducted in two different ways: first, as a binary choice task and then, with a Confidence scale in which participants were asked to rank their trust in their binary response in a scale from 1 to 3. The Confidence scale was presented on screen to remind participants that 1 meant "Very insecure" and 3 "Very confident", and it was programmed to translate every registered response into a larger scale (with values from 1 to 6) depending on their previous binary response. According to this new scale, extreme values 1 and 6 referred a greater confidence in their previous response (with 1 being "Very sure it was a Noise trial" and 6 being "Very sure it was a Signal trial"), while values in the middle could serve as indicators of greater uncertainty (so that 2 and 3 would mean "I'm very insecure about the previous trial being a Noise trial" or ""I'm very insecure about the previous trial being a Noise trial). This transformation procedure has been used in recognition memory studies where the mirror effect had been reported (Glanzer y Adams, **?**), with the intention of making the rating task easier for participants, while guaranteeing a match between their binary response and the value assigned in the full scale.

The experiment consisted of 640 trials, with each one of these following the next structure:

- **First part: Binary choice Yes/No**

  At the beginning of each trial, the pair of central circles that participants had to compare appeared on screen along the reminding legends "Do the central circles have the same size?" on top of the screen and the answer keys on the bottom "S = Yes, N = No". After 1.5 seconds, the stimuli disappeared from the screen to prevent habituation, while the reminding legends remained until a response was registered.

- **Second part: Confidence Scale**

  Once participants had registered their binary response to the Yes/No task, the Confidence scale was shown below the instructional question "How certain are you about your response?". The Confidence scale shown contained numbers 1, 2 and 3 along with the legends "Very insecure", "More or less sure" and "Very confident".

- **Third part: "Space-bar to continue"**

  Right after a response to the Confidence scale was registered, a third and final screen was shown to indicate the end of the trial with the legend "Press the space-bar to move on to the next trial" . This final screen was included with the explicit intention of avoiding participants' fatigue by giving them the chance to take a break before continuing to be exposed to more Ebbinghaus figures.

At the end of the 640 trials, participants were shown the total number of right and wrong responses registered (which was included just to provide participants' with a sense of ending). We never told participants the purpose of the experiment and more importantly, participants were never told anything about the presence of two distinct classes of stimuli coexisting within the same task.

3 Results

Both experiments shown evidence of the patterns of response identified as part of the mirror effect in more than three fourths of the participants. In Experiment 1, where a total of 20 participants did the task, 17 showed the mirror effect pattern of response in the binary task and 18 in the Confidence scale. In Experiment 2, 18 participants showed evidence of the mirror effect patterns of response in both the binary task and the confidence rate task. These proportions shown to be statistically significant when analyzed with a simple binomial test.

It is important to note that prior to the conduction of any kind of formal data analysis, we examined the responses registered for every participant to make sure that the data obtained was consistent with the task we proposed. While reviewing participants' responses registered in every consecutive trial we found that Participant No.1 in Experiment 2 spent the very first eighty trials responding to the answer key for "No", (view Fig **??**). This participant was excluded from any further data analysis.

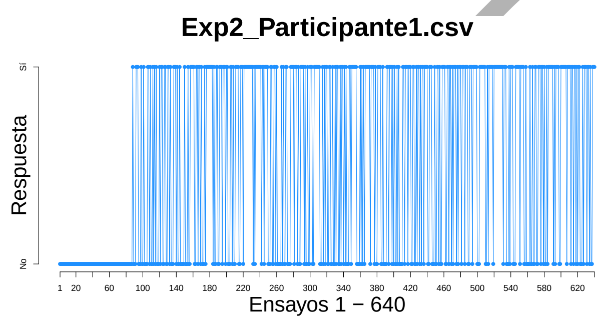


Figure 6: Yes/No responses registered on each of the 640 trials by participant No.1 in Experiment 2. A clear pattern can be observed during the first 80 trials, which suggest that participant's responses were not made according to the instructions or the stimulus presented.

To make the results obtained in the present study comparable to what has been reported in the recognition memory study where the mirror effect has been exposed, we conducted a step by step replication of the data analysis reported by Glanzer and Adams, (**?**);

1. **Comparing d' across classes of stimuli.** This initial test is considered mandatory in order to guarantee that classes A and B are indeed different in terms of how

   A t-test was conducted to compare the estimations for d' obtained for each class of stimuli, where statistically significant differences were found in the direction that we expected according to our experimental design.

2. Comparing the Hit and False Alarm rates obtained in each class of stimuli

   Two separate t-tests were conducted to compare separately the Hit rates and the False alarm rates registered between each class of stimuli. Statistically significant differences were found in the same direction as reported in the recognition memory literature.

3. Comparing the mean Confidence ratings registered in each class of stimuli

> Two separate t-tests were conducted to compare separately the mean Confidence rates registered for the Signal and Noise trials in each class of stimuli. Statistically significant differences were found in the same direction as reported in the recognition memory literature.

4 Bayesian Cognitive Modeling

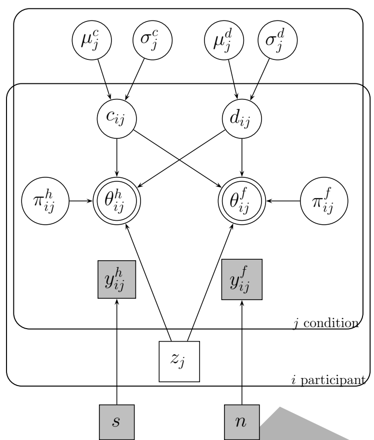
The present study aims to encourage the application of Bayesian methods to Cognitive Sciences by illustrating the advantages it can provide to the assessment of the core assumptions made by Cognitive Models in terms of their reflection on the data gathered in an experimental task, as well as for the righteous representation and manipulation of individual differences.

4.1 Contaminant Models

One of the many advantages of the application of Bayesian methods to the development of Cognitive Models is that it allows us to test the usually accepted assumption that all data obtained in an experimental task is suitable to be explained as the result of the high-level mechanisms that are assumed by a certain statistical model that is being treated as a cognitive model. The development of latent class mixture models, (Chávez, M, Villalobos, E., Baroja, J.L & Bouzas, A., **?**; Velázquez, C., Villarreal M., & Bouzas, A., **?**).

In the present study, we developed two different Contaminant models. The first one focused on looking for contaminant data at the level of the overall responses registered during the whole experiment , while the second one was more focused on every single response registered by each participant during the 640 trials. In other words, the first general model allows us to make sure that all the observed patterns of Hits and False alarms rates can be explained within the SDT framework, while the second, more specific approach allow us to detect participants whose patterns of response suggest that there was no independence among trials, (just like Participant shown in the Figure **??**, who kept responding to the "No" key for the first 80 trials). It is important to know that given the way the experiment was programmed, all different types of stimuli were presented at random (both in terms of the "A" and "B" class, and in terms of Signal and Noise stimuli), we would expect participants to respond to each key with an equal probability among trials.

The graphical model that illustrates the first general Contaminant model is shown in Figure **??**. For those who may not be familiar with the use of Graphical models to represent the inner structure assumed by any statistical model in terms of how data is generated from a certain, underlying process, when parameters proposed to measure specific aspects or processes interact, we strongly recommend to read any of the following materials (Lee and Wagenmakers, **?**; Lee, **?**,).



$$\mu_j^c \sim \text{Gaussian}(0, 0.7)$$
$$\mu_j^d \sim \text{Gaussian}(0, 1)_{T(0,\infty)}$$
$$\sigma_j^c, \sigma_j^d \sim \text{Uniform}(0, 4)$$
$$d_{ij} \sim \text{Gaussian}(\mu_j^d, \sigma_j^d)$$
$$c_{ij} \sim \text{Gaussian}(\mu_j^c, \sigma_j^c)$$
$$\pi_{ij}^h, \pi_i^f \sim \text{Beta}(1, 1)$$
$$\theta_{ij}^h \leftarrow \begin{cases} \pi_{ij}^h & \text{if } z_j = 0 \\ \phi(\frac{1}{2}d_{ij} - c_{ij}) & \text{if } z_j = 1 \end{cases}$$
$$\theta_{ij}^f \leftarrow \begin{cases} \pi_{ij}^f & \text{if } z_j = 0 \\ \phi(-\frac{1}{2}d_{ij} - c_{ij}) & \text{if } z_j = 1 \end{cases}$$
$$z_j \sim \text{Bernoulli}(0.5)$$
$$y_{ij}^h \sim \text{Binomial}(\theta_{ij}^h, s)$$
$$h_{ij}^f \sim \text{Binomial}(\theta_{ij}^f, s)$$

Figure 7: Latent-Mixture model to detect contaminant participants withn each sample, based on their overall performance during the task.

In our first Contaminant model, every participant i in the class of stimuli j, the total number of Hits (y_{ij}^h) and False alarms (y_{ij}^f) made along the total number of Signal (s) and Noise (n trials, was attributed to their own individual probabilities of making a Hit (θ_{ij}^h) or a False Alarm (θ_{ij}^f). This underlying probability could arise from two different sources, depending on an individual dicotomic parameter z_j. For participants whose z had a value of 0, these probabilities would be equal to a "random" parameter (π_{ij}^h, and π_{ij}^f, respectively), which had a uniform prior distribution from 0 to 1. On the other hand, participants whose z value was 1, had these probabilities drawn from a SDT process described by the interaction of the discriminability parameter d_{ij} and the bias parameter c_{ij}, which were drawn from two distinct group-level distribution for each class of stimuli.

According to the core assumptions held under a Signal Detection Theory framework, the total number of Hits registered by any participant on any signal-detection task must be greater than the total number of False Alarms registered in the same task by the same participant. This general pattern is implied by the core assumption that signal stimuli, in general, have "more" of whatever is contained in the evidence axis and to what participants are expected to be responding, than the noise stimuli. To have more False Alarms being registered than Hits, would contradict this assumption by suggesting a particular scenario where the Noise distribution is located to the right of the Signal distribution. The main goal of this first contaminant model, was to

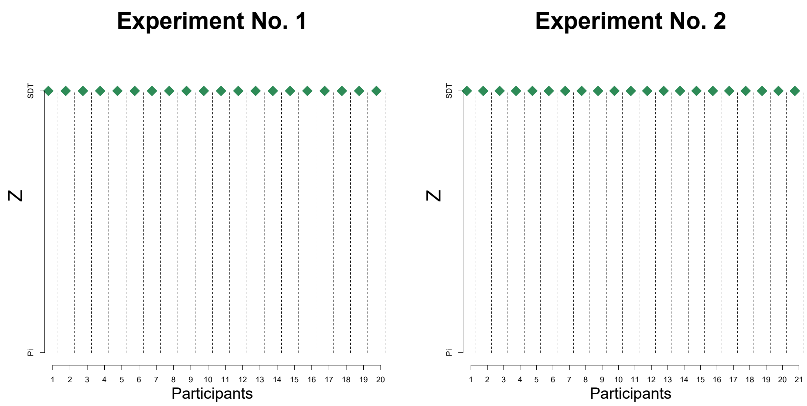


Figure 8: Individual estimations for every participant in experiments 1 and 2, on the value of z, the dicotomic latent parameter controlling whether the probabilities of making a Hit and a False Alarm could be described by the SDT framework, or if they could be interpreted as regular, random probabilities

make sure that the patterns of Hits and False Alarms recorded by each participant on each class of stimuli, satisfied this general rule.

Results obtained in this first contaminant model are shown in Figure **??**. In both experiments 1 and 2, all participants shown values of 1 on the latent dicotomic parameter z, thus demonstrating that data collected was suitable to be described by a SDT framework.

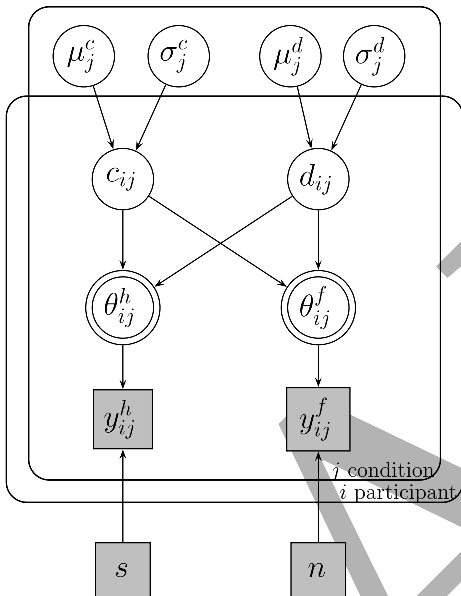
The second Contaminant model is shown in Figure **??**,

4.2 Hierarchical SDT Model

After checking for the presence of contaminant data, we proceeded to work on the Bayesian Cognitive modeling of the data that had already demonstrated . Our initial move was to conduct a simple Hierarchical extension of the Signal Detection Model, as presented in the Graphical Model on Figure **??**.

This model constitutes the base for the great majority of the forthcoming Bayesian models, from its representation of the underlying processes assumed by SDT (which was originally developed by Lee and Wagenmakers, **?**), to the specific priors used to reflect assumptions derived from our experimental task, concerning parameters c_{ij} and d_{ij}.

According to our Hierarchical SDT model, the total number of Hits (y^h_{ij}) and False Alarms (y^f_{ij}) registered by every participant i across each class of



$$\mu_j^d \sim \text{Gaussian}(0,1)_{T(0,\infty)}$$
$$\mu_j^c \sim \text{Gaussian}(0,0.7)$$
$$\sigma_j^c, \sigma_j^d \sim \text{Uniform}(0,4)$$

$$d_{ij} \sim \text{Gaussian}(\mu_j^d, \sigma_j^d)$$
$$c_{ij} \sim \text{Gaussian}(\mu_j^d, \sigma_j^c)$$

$$\theta_{ij}^h = \phi(\tfrac{1}{2}d_{ij} - c_{ij})$$
$$\theta_{ij}^f = \phi(-\tfrac{1}{2}d_{ij} - c_{ij})$$

$$y_{ij}^h \sim \text{Binomial}(\theta_{ij}^h, s)$$
$$y_{ij}^f \sim \text{Binomial}(\theta_{ij}^f, n)$$

Figure 9: Graphical model for the Hierarchical extension of the SDT model. This model assumes that both individual d' and individual c values are drawn from a group-level normal distribution.

stimuli j out of the total number of signal (s) and noise (n) stimuli presented during the task, could be described as the result of a binomial process with the probability of making a Hit or a False Alarm on each class of stimuli being captured by θ_{ij}^h and θ_{ij}^h, respectively. The SDT framework is captured within these models in the way that these theta parameters are derived from the interaction of d_{ij} (the parameter measuring discriminability) and c_{ij} (the parameter quantifying subjects' bias). The general idea is that each theta is equal to the cumulative density of its corresponding distribution (signal or noise), which is assumed to be normal, above the criterion, (Lee Wagenmakers, **?**; Gescheider, **?**). In this hierarchical SDT model designed to merely explore our data under an SDT framework, both parameters c and d were assumed to be drawn from a group-level normal distribution, defined by their own means and standard deviation.

Results from the Hierarchical SDT modeling can be seen in Figure **??**, where both the joint and the marginal posterior estimations for d and c, on Experi-

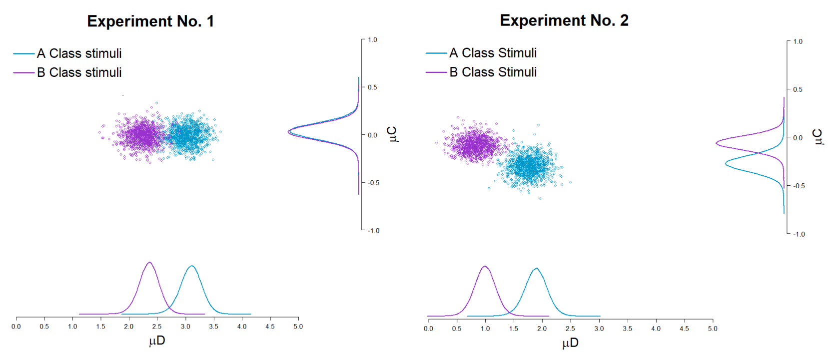


Figure 10: Joint and marginal posteriors for the mean and

ments 1 and 2, can be seen. From the marginal posterior distributions drawn for d across each class of stimuli,

4.3 Comparing d' across classes of stimuli

In order to ascertain the validity of our experimental design, where we designed two different classes of stimuli according to the literature on the Ebbinghaus illusion, we developed a Bayesian model

4.4 Comparing c across classes of stimuli

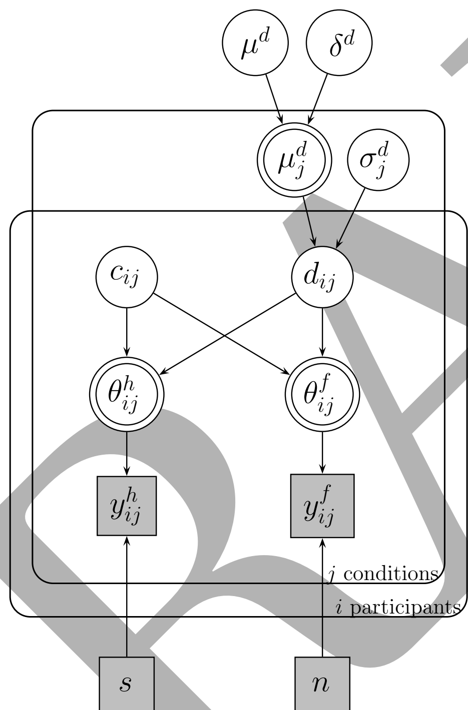
4.5

4.6

5 Discussion

6 Conclusion

7 Acknowledgements



$$\mu^d \sim \text{Gaussian}(0,1)_{T(0,\infty)}$$
$$\delta^d \sim \text{Gaussian}(0,1)_{T(0,\infty)}$$

$$\mu_A^d = \mu^d + \tfrac{\delta}{2}$$
$$\mu_B^d = \mu^d - \tfrac{\delta}{2}$$
$$\sigma_j^d \sim \text{Uniform}(0,5)$$

$$d_{ij} \sim \text{Gaussian}(\mu_j^d, \sigma_j^d)$$
$$c_{ij} \sim \text{Gaussian}(0, 0.7)$$

$$\theta_{ij}^h = \phi(\tfrac{1}{2}d_{ij} - c_{ij})$$
$$\theta_{ij}^f = \phi(-\tfrac{1}{2}d_{ij} - c_{ij})$$

$$y_{ij}^h \sim \text{Binomial}(\theta_{ij}^h, s)$$
$$y_{ij}^f \sim \text{Binomial}(\theta_{ij}^f, n)$$

Figure 11: A boat.

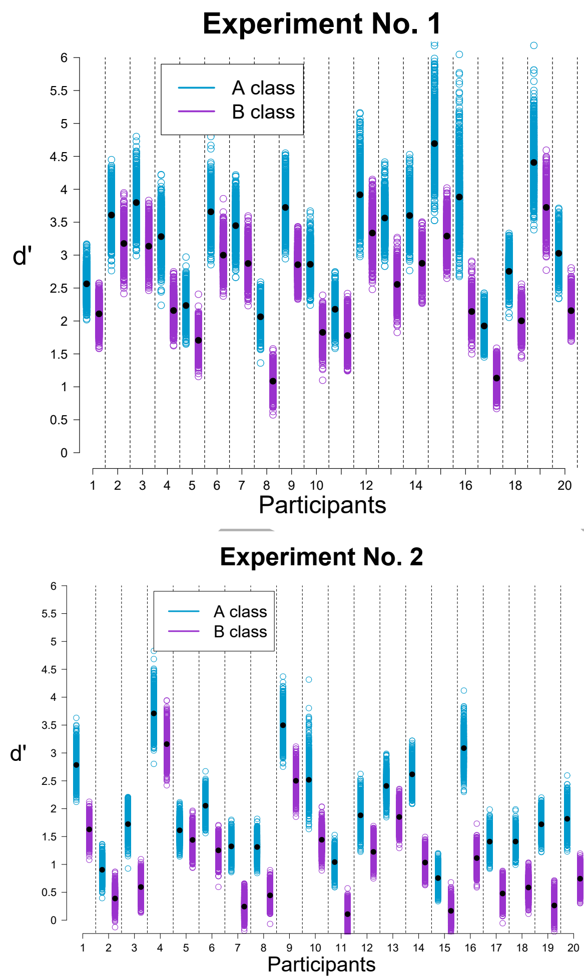


Figure 12: Densities

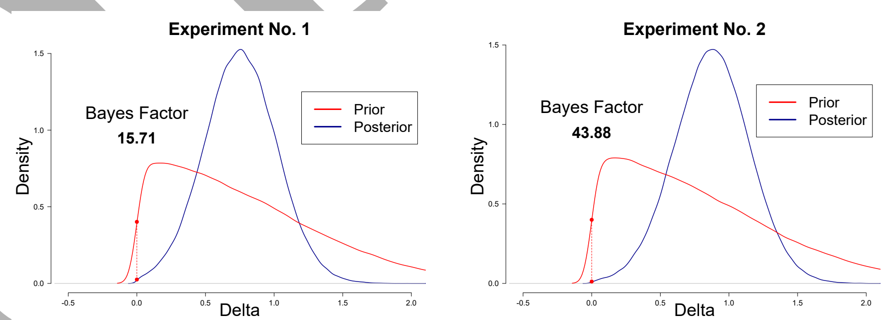
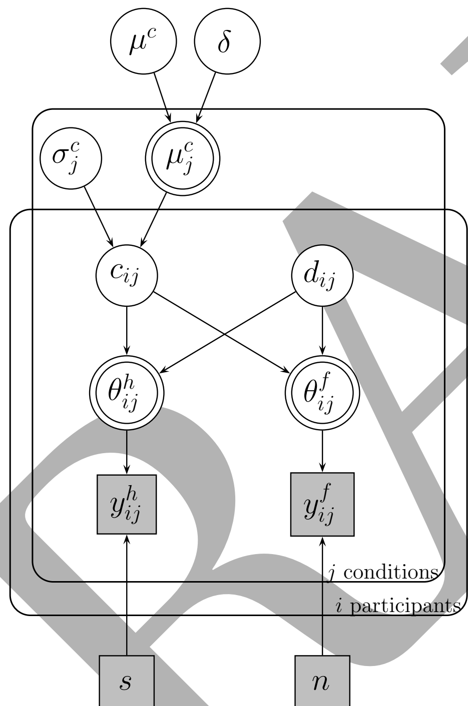


Figure 13: Bayes Factor



$$\mu^c \sim \text{Gaussian}(0, 0.7)$$
$$\delta \sim \text{Gaussian}(0, 1)$$

$$\mu_A^c \leftarrow \mu^c + \tfrac{\delta}{2}$$
$$\mu_B^c \leftarrow \mu^c - \tfrac{\delta}{2}$$
$$\sigma_j^c \sim \text{Uniform}(0, 5)$$

$$c_{ij} \sim \text{Gaussian}(\mu_j^c, \sigma_j^c)$$
$$d_{ij} \sim \text{Gaussian}(0, 1)_{T(0,\infty)}$$

$$\theta_{ij}^h \leftarrow \phi(\tfrac{1}{2}d_{ij} - c_{ij})$$
$$\theta_{ij}^f \leftarrow \phi(-\tfrac{1}{2}d_{ij} - c_{ij})$$

$$y_{ij}^h \sim \text{Binomial}(\theta_{ij}^h, s)$$
$$y_{ij}^f \sim \text{Binomial}(\theta_{ij}^f, n)$$

Figure 14: A boat.